Implications of Model Choice in Lévy-driven Financial Markets

Maren Diane Schmeck

DISSERTATION PRESENTED FOR THE DEGREE OF PHILOSOPHIAE DOCTOR

CENTRE OF MATHEMATICS FOR APPLICATIONS

UNIVERSITY OF OSLO

2012

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 1275

ISSN 1501-7710

Cover: Inger Sandved Anfinsen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika publishing.
The thesis is produced by Akademika publishing merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.

SUMMARY

In the process of modeling, the agent has to make a number of decisions. Therefore two agents can set up different models. This thesis compares the influence of different model choices to stochastic control theory and option pricing. On the one hand it quantifies analytically error bounds in terms of key parameters. On the other hand it checks the implications of model choices empirically.

Lévy processes are popular in financial modeling since they are able to explain many of the stylized facts of asset prices. In particular, some processes like the normal inverse Gaussian (NIG) or the hyperbolic Lévy process have become particularly relevant since they are able to capture the return distribution of most asset prices. These Lévy processes are pure-jump, and therefore give distinctively different paths of the asset prices compared to a Brownian motion with continuous paths. In empirical analysis of financial price data, one may detect big jumps, however, the small jumps are very hard to separate from the observations of a Brownian motion. Thus, it is not a simple task to decide whether a Lévy process with jumps or a Brownian motion is governing the small variations in a stock price, say. First Merton's portfolio optimization problem is considered. We aim for a mathematical quantification of the difference of the optimal investment strategies and find error bounds that are proportional to the variation of the small jumps. Then option pricing is considered, where there are two underlying assets that are dependent. Here the error bounds turn out to be of the same type as in Article 1.

The second part of the thesis considers the pricing of options on forwards in energy markets. Many models for the electricity spot price divide the price evolutions into a short-term and a long-term component. Electricity markets are well-known for their large price variations and rare, big spikes, which are captured by the short-term component. We examine the influence of the short-term and long-term factor on the spot and prove that the short-term factor is insignificant for pricing options in many relevant cases.

The electricity spot is not a tradable asset and the no-arbitrage argumentation that is based on cost and carry strategies cannot be applied. Therefore one can use different measures to price the futures and the option. The goal of this paper is to examine empirically wether or not the traded options are priced under the same pricing measure as the futures. Before doing so, we need to specify a model for spot and futures and fit it to data. The results indicate that one should use a different measure to price the option.

ACKNOWLEDGMENTS

I am grateful to many people for their support and help during my work on the thesis. First of all I would like to thank my principal supervisor Fred Espen Benth for his guidance, support and inspiring collaboration. This thesis would not exist without him. I also would like to thank my subsidiary supervisor Erik Bølviken as well as my co-authors Giulia Di Nunno and Asma Khedher for sharing their experience and expertise. Furthermore I am grateful to Jukka Lempa for his fruitful comments and discussions, and to Hanspeter Schmidli, who gave me the time to prepare for the defence. Also I would like to thank Daniel Bauer and Rocio Sotomayor. They gave me the opportunity for a research visit at the J. Mack Robinson College of Business, Atlanta, and welcomed me with open arms.
I considered myself very lucky for beeing a member of the Centre of Mathemtics for Applications (CMA). I would like to thank the administration, Ragnar Winther, Helge Galdal and Aslaug Kleppe Lyngra, for providing such an excellent and well-administrated working environment and my collegues and fellow Phd-students for creating such a friendly and inspiring atmosphere at the institute.
A special thanks goes to my friends, who accompanied me during the last three years through all the ups and downs. Last but not least, I am grateful to Vegard for his solidarity while finalising the thesis, and my parents and siblings for their unlimited patience and dedication.

Blindern, December 2012, Maren Diane Schmeck

CONTENTS

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

This thesis consists of four articles that deal with different applications of mathematical finance. They have in common that they compare the performance of different models. The first three articles quantify analytically their fit in terms of key parameters, while the fourth article examines different models empirically.

1.1 Model choice and implications

Consider a filtered complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ and a stochastic process $X_\alpha = (X_\alpha(t))_{t\geq 0}$, that depends on a parameter $\alpha \in \mathbb{R}$ and that describes the evolution of a risky asset in time. Suppose that the noise of X_α comes from a Lévy process with characteristic triple (a, σ, ν). Fix a parameter $\alpha_0 \neq \alpha$, $\alpha_0 \in \mathbb{R}$, and think of

$$X_\alpha \qquad \text{and} \qquad X_{\alpha_0}$$

as two different models for the risky asset. At each point of time $t \geq 0$ they converge to each other in distribution:

$$X_\alpha(t) \to X_{\alpha_0}(t) , \qquad (\alpha \to \alpha_0) .$$

Furthermore, consider an application with underlying process X_α that results in a number of interest $y_\alpha \in \mathbb{R}$ which is specified through a functional F:

$$y_\alpha = F(\alpha; a, \sigma, \nu) \in \mathbb{R} . \tag{1.1.1}$$

The application includes an expectation, such that y_α is determined as a function of the characteristic triple of the underlying Lévy process. We are now interested in a quantification of the error of the form

$$|y_\alpha - y_{\alpha_0}| \leq |E(\alpha, \alpha_0)| , \qquad (\alpha \to \alpha_0) ,$$

for another function E giving the convergence rate and with $E(\alpha, \alpha_0) \to 0$, $(\alpha \to \alpha_0)$.

In the first two articles the models X_{α_0} and X_α are based on the same method. Here, the models differ in how the small variations in the risky asset behave, by infinitely many small jumps or a continuous Brownian motion. That is, the risky asset is driven by a pure jump Lévy process with infinite activity

$$L(t) = \int_0^t \int_{|z|<1} z\tilde{N}(ds, dz) + \int_0^t \int_{|z|\geq 1} zN(ds, dz) \ , \tag{1.1.2}$$

or a Lévy process of the form

$$L_\epsilon(t) = \sigma(\epsilon)W(t) + \int_0^t \int_{\epsilon<|z|<1} z\tilde{N}(ds, dz) + \int_0^t \int_{|z|\geq 1} zN(ds, dz) \, , \tag{1.1.3}$$

where W is a Brownian motion (independent of L) and

$$\sigma^2(\epsilon) := \int_{|z|<\epsilon} z^2 \nu(dz)$$

is the variance of the small jumps. To get from (1.1.2) to (1.1.3), neglect the jumps with absolute size smaller than $\epsilon > 0$ and approximate them by a scaled Brownian motion. Here, the parameter is the level of truncation $\alpha = \epsilon$, where $\alpha_0 = 0$ means that there is no truncation. The process L corresponds then to the model X_{α_0} and L_ϵ corresponds to X_α. In fact, all noise can be described using the components of one underlying Lévy process with characteristic triple $(0, 1, \nu)$.

Article [1] considers Merton's portfolio optimization problem, where one is interested in the fraction of the wealth of an investor, that she should invest into the risky asset. As driving processes, she can choose (1.1.2) or (1.1.3). The optimal controls

$$y_\alpha = \pi^*_\epsilon \qquad \text{and} \qquad y_{\alpha_0} = \pi^*$$

are constant in time and the solution of an integral equation, that can be specified through a function g of the form

$$g(\pi) := \int_{\mathbb{R}\setminus\{0\}} [1 + \pi(e^z - 1)]^{\gamma-1}(e^z - 1) - (e^z - 1)\nu(dz) \, .$$

We prove that g is invertible such that we can set $F = g^{-1}$. Then we specify the error as

$$|\pi^*_\epsilon - \pi^*| \leq C\sigma^2(\epsilon) \, , \tag{1.1.4}$$

and conclude the stability of Merton's portfolio optimization problem with respect to model choice. For a small ϵ it does not matter too much which model one chooses.

In Article [2] we have a bivariate asset price process, where we allow for dependency between the two stocks. Therefore, we have a bivariate Lévy process as driving noise. We perform an approximation of the small jumps analogous to (1.1.2) and (1.1.3). Then we consider options with the bivariate asset dynamics and its approximation as underlying. They have the option prices

$$y_{\alpha_0} = C \qquad \text{and} \qquad y_\alpha = C_\epsilon \ .$$

We write the option prices as an expression based on a Fourier-transform, which results in a function F as in (1.1.1) and find an error of the same type as in (1.1.4). Again we conclude that the choice of model does not matter. We consider spread options, where one bets on the difference of two risky assets.

Article [3] discusses pricing and hedging of options in energy markets. Big spikes can generally appear in the spot price data, that return quickly back to the normal price level. We choose a two-factor model for the spot, and derive the following dynamics for the price $X_\alpha = f_\beta$ of a future with delivery at time T:

$$\frac{df_\beta(t,T)}{f_\beta(t-,T)} = \sigma\, dB(t) + \int_{\mathbb{R}} \left\{ \exp\left(z\mathrm{e}^{-\beta(T-t)} \right) - 1 \right\} \widetilde{N}(dz,dt)\,.$$

The parameter β describes the speed of mean reversion of the spikes in the spot. For $\alpha_0 = \infty$, we have a geometric Brownian motion

$$\frac{df(t,T)}{f(t,T)} = \sigma dB(t)\,.$$

One can think of a speed of mean reversion that is infinitely fast as corresponding to the non-existence of spikes. We are interested in the prices

$$y_\alpha = C_\beta \qquad \text{and} \qquad y_{\alpha_0} = C$$

of an European call option with the futures as underlying and exercise time $\tau \leq T$. For C, the Black-76 formula gives an explicit expression and we derive an expression of similar type for C_β. We can quantify the error between C and C_β as

$$|C_\beta - C| \leq c\mathrm{e}^{-\beta(T-\tau)}\,,$$

for some constant c. We conclude that if $\beta(T-\tau)$ is big, then extreme spikes do not influence the option price and the Black-76 formula is a good approximation. We prove the same result also for the quadratic hedging component.

Model choice in Article [4] is about choosing an appropriate measure to price options on futures in energy markets. The lack of no-arbitrage conditions allows one to detach the pricing measures, that are used to price the future and the option. In an empirical manner we address the question of whether the market uses the same measure to price options as to price futures, or not. Therefore, we compare quoted option prices with simulated prices based on a two factor model for the spot, and furthermore with Black-76 prices. The results indicate that the market might use a different measure.

1.2 The Brownian approximation of small jumps

In the classical theory of financial mathematics the dynamics of the risky asset is driven by a Brownian motion. In this setting, Merton [49] proved in 1969 that a risk averse investor will place a constant proportion of her total wealth in the risky asset. In 1973 Black and Scholes [22] published their famous theory for option pricing where the risky asset follows a geometric Brownian motion, and in 1976 Black [21] extended this theory for options with a future as underlying. Since then these results where extended to the case where the driving noise is a Lévy process and beyond that to more general stochastic processes. See for example Cont and Tankov [31] for financial markets including jump processes and see Øksendahl and Sulem [58] and Benth et al. [13, 14] for stochastic control theory in a jump diffusion setting.

One reason for the popularity of Lévy models in finance might be, that there are distributions connected to pure jump Lévy processes, that are able to describe the (log-) returns of financial assets in different markets better than the normal distribution. Especially the normal inverse Gaussian distribution (NIG), belonging to the class of hyperbolic distributions, captures many stylized facts of data well and convinces furthermore through its analytical tractability. A normal inverse Gaussian process is a Lévy process L, where $L(1)$ is NIG distributed. It has an explicit expression for the characteristic function, stays a NIG after an Esscher transform and the behaviour of the variation of the small jumps is known. In this thesis there are many examples and numerical illustrations using the NIG distribution (see the examples in Section 2.4.4 and 4.3 and Figure 2.2, 4.1 and 5.3). For an introduction to the hyperbolic distribution, see Eberlein and Keller [35] and to the normal inverse Gaussian distribution and the corresponding NIG process, see Barndorff-Nielsen [6] and Rydberg [51].

After the Lévy-Itô decomposition, a Lévy process can be written as a linear combination of a drift, a scaled Brownian motion and a pure jump process with stationary and independent increments. Within the class of Lévy processes, the Brownian and jump component are independent, and therefore the two parts do not mix in applications either, but can be identified as separate expressions. This makes it comfortable to vary within the class of Lévy processes, to add a Brownian term or a jump component to the setting and observe the impact. In a way, that is what we do in the articles [1] and [2].

Consider a Lévy process L with infinite activity and Lévy-Itô decomposition

$$L(t) = \int_0^t \int_{|z|<1} z\tilde{N}(ds, dz) + \int_0^t \int_{|z|\geq 1} zN(ds, dz) \,. \tag{1.2.1}$$

Consider not only one, but two approximations of L. First, truncate the jumps with absolute value smaller than $\epsilon > 0$

$$L_{N,\epsilon}(t) = \int_0^t \int_{\epsilon\leq|z|<1} z\tilde{N}(ds, dz) + \int_0^t \int_{|z|\geq 1} zN(ds, dz) \,. \tag{1.2.2}$$

Then, in a second step, replace them by Brownian motion W (independent of L)

$$L_{W,\epsilon}(t) = \sigma(\epsilon)W(t) + \int_0^t \int_{\epsilon\leq|z|<1} z\tilde{N}(ds, dz) + \int_0^t \int_{|z|\geq 1} zN(ds, dz) \,, \tag{1.2.3}$$

where W is scaled by $\sigma(\epsilon)$ with

$$\sigma^2(\epsilon) := \int_{|z|<\epsilon} z^2\nu(dz) \,. \tag{1.2.4}$$

The quality of the approximation (1.2.3) is dependent on the speed at which $\sigma^2(\epsilon)$ converges to 0 for $\epsilon \to 0$. The Brownian approximation (1.2.3) is based on results from Asmussen and Rosinski [5]. They prove that the small jumps of a Lévy process with infinite activity are indeed approximately Gaussian, that is that they have a central-limit type behavior towards a Brownian motion, if the condition

$$\frac{\sigma(\epsilon)}{\epsilon} \to \infty \qquad \text{for } \epsilon \to 0 \tag{1.2.5}$$

is fulfilled. For the normal inverse Gaussian process the Condition (1.2.5) is valid, as

$$\sigma(\epsilon) \sim \epsilon^{\frac{1}{2}} \qquad \text{for } \epsilon \to 0 \ . \tag{1.2.6}$$

An example where Condition (1.2.5) does not hold is the Gamma process, as it is

$$\sigma(\epsilon) \sim \epsilon \qquad \text{for } \epsilon \to 0$$

(see Asmussen and Rosinski [5]). These results where motivated from Rydberg [51], who approximated intuitively the small variations of a NIG-process through a Brownian motion. Cohen and Rosinski [30] showed that the convergence result holds as well in the multivariate case. Based on (1.2.1) - (1.2.3) Cont and Tankov [31] consider methods for simulation of the processes as well as option prices and specify convergence rates depending on $\sigma^2(\epsilon)$. Neglecting the small jumps leads to the so-called Poisson-approximation, or series representation. Replacing the small jumps by a Brownian motion can increase the accuracy of the Poisson-approximation. This is dependent on the growth of the Lévy measure around 0. They point out that (1.2.3) does not always lead to an improvement compared to (1.2.2) as for example in the case of the Gamma process.

Merton's portfolio selection problem has gained a lot of attention in the scientific literature over the years. Merton [49] proved that a risk averse investor will place a constant proportion of her total wealth in risky assets. Since then, it has been examined under modifications in many directions. Benth, Karlsen and Reikvam [13] studied Merton's problem when the risky asset price dynamics is given by an exponential pure-jump Lévy process. As it turns out, it is also in this case optimal to invest a constant proportion of the investor's wealth in the risky asset. However, the proportion is given in terms of the solution of an integral equation involving the excess return of the asset and the characteristics of the jump process.

One can think of the Brownian approximation also in the following way. Empirically it is very hard, if not impossible to decide if the observed paths of the assets are continuous or driven by infinitely many small jumps. Therefore one could think of two investors with the same risk aversion, one believing that the small variations in the asset price dynamics are continuous, the other believing in small jumps. They coincide about the existence and modeling of big jumps. Both state the corresponding optimal investment problem, solve it and find their optimal control. It might be that the optimal investment strategies coincide, and are only stated in different terms, in the language of jump- or continuous processes. We aim therefore for a mathematical quantification of the difference of the investment strategies, as well as the value functions and wealth processes. This is the main interest of Article [1].

Benth et al. [9, 10, 11] study comprehensively the stability of option prices and their deltas with respect to model choice in one-dimensional jump diffusion models. They prove $\mathcal{L}^2$ convergence of the approximated Lévy processes as well as solutions of jump diffusions. Furthermore, they consider the stability of option prices under a change of measure, where the measure depends on the choice of model and study applications to stochastic volatility

models. Khedher [44] extends the results of the stability of option prices to a bivariate setting with independent underlying. Article [2] continues this study by including dependency into the model, as the price of an option on a multivariate underlying will be dependent on the correlation between the assets. It furthermore concentrates on the convergence rates and gives improved rates in the case when one approximates by a Brownian motion. In particular, we consider Margrabe's formula for spread options. Spread options bet on the difference of two risky assets, such that the value of the option will be influenced by the dependency structure of the underlying. Margrabe's formula for spread options (see Margrabe [48]) is very popular in literature, see for example Carmona and Durrleman [28], Cheang and Chiarella [29], Eberlein et al. [36] and Benth et al. [12]. The popular formula transforms the spread option on a bivariate underlying into a plain vanilla option on a one dimensional underlying.
Another approach for pricing spread options can be found in Borovkova et al. [25]. They derive closed form approximations for pricing spread options by approximating the spread distribution using a generalized family of log-normal distributions.

In the next two sections it follows a summary of Article [1] and Article [2].

1.2.1 Summary of Article 1: "Stability of Merton's Portfolio Optimization Problem for Lévy Models"

We start the article with a review of Merton's portfolio optimization problem. We assume that the stock price follows a geometric jump diffusion which is driven by a pure jump Lévy process L with infinite activity as in (1.2.1). To ensure that the first two moments of the stock price process are finite, we impose an integral condition on the Lévy process exponential integrable up to order 2. The wealth process is again a geometric jump diffusion. The investor optimises her utility coming from consumption, and she has a power utility function of the form $U(x) = x^\gamma/\gamma$ for a risk aversion parameter $\gamma \in (0,1)$. Denote by π the fraction of her wealth that she invests in the stock . Letting $\delta > 0$ be a constant discount rate, the value function is defined by

$$V(x) = \sup_{c,\pi \in \mathcal{A}_x} \mathbb{E}^x \Big[\int_0^\infty e^{-\delta t} [\frac{c_t^\gamma}{\gamma}] dt \Big] .$$

Here, we restrict ourselves to admissible controls constrained by $[0,1]$, which means that we do not allow short selling of stocks or borrowing money to invest more than our wealth in the stock. With some additional effort it is possible to extend the theory to more general bounds.

Benth et al. [13] show that the optimal investment strategy turns out to be a constant π^* solving implicitly the integral equation

$$g(\pi) = r - \hat{\mu} \,,$$

where

$$g(\pi) := \int_{\mathbb{R}\setminus\{0\}} [1 + \pi(e^z - 1)]^{\gamma-1}(e^z - 1) - (e^z - 1)\nu(dz) \,.$$

The optimal consumption is given as a constant rate of the wealth, see (2.2.6).

Then we examine Merton's portfolio problem when we approximate the driving Lévy process L in the stock price dynamics. In particular, we consider the two approximations where the small jumps of L are neglected and where we substitute the small jumps by a scaled Brownian motion. These approximations will lead to different HJB-equations, and thus to different controls and value functions. First, consider the approximation $L_{N,\epsilon}$ in (1.2.2) of L which neglects jumps smaller than ϵ. The neglection of the small jumps influences the Levy measure and so indirectly all terms which include a Lévy integral, as the drift of the stock price dynamics, the optimal controls and eventually the value function. Tracing through the derivation of Benth et al. [13] leads to the following (reformulated) integral equation for the optimal control $\pi^*_{N,\epsilon}$

$$g(\pi_{N,\epsilon}) = r - \hat{\mu} + \int_{|z|<\epsilon} h(\pi_{N,\epsilon}, z)\, \nu(dz) \,, \tag{1.2.7}$$

where for $|z| < 1$

$$h(\pi_{N,\epsilon}, z) := [1 + \pi_{N,\epsilon}(e^z - 1)]^{\gamma-1}(e^z - 1) - z \,.$$

Now, consider the approximation $L_{W,\epsilon}$ as in (1.2.3). It gives the following integral equation for the optimal investment strategy $\pi^*_{W,\epsilon}$

$$g(\pi_{W,\epsilon}) = r - \hat{\mu} + \sigma^2(\epsilon)\Big((1-\gamma)\pi_{W,\epsilon} - \frac{1}{2}\Big) + \int_{|z|<\epsilon} h(\pi_{W,\epsilon}, z)\, \nu(dz) \,. \tag{1.2.8}$$

Additionally to the term $\int_{|z|<\epsilon} h(\pi_{W,\epsilon}, z)\, \nu(dz)$, that also appeared in (1.2.7), we have a term with $\sigma^2(\epsilon)$ on the right hand side of (1.2.8).

Lemma 2.2.1 gives conditions which ensure the existence and uniqueness of an optimal portfolio investment strategy $\pi^* \in [0, 1]$ with underlying process L. We show that the conditions in Lemma 2.2.1 are sufficient for the existence and uniqueness of an optimal strategy $\pi^*_{N,\epsilon} \in [0, 1]$ and $\pi^*_{W,\epsilon} \in [0, 1]$.

We move on with the stability analysis of the optimal controls. Proposition 2.4.1 states that the approximative investment strategy $\pi^*_{N,\epsilon}$ converges to π^* as $\epsilon \to 0$ with a convergence rate proportional to the variance of the small jumps $\sigma^2(\epsilon)$:

$$|\pi^*_{N,\epsilon} - \pi^*| \le C_N \sigma^2(\epsilon) \,,$$

where $C_N > 0$ is a constant independent of ϵ. In the proof we show that g^{-1} exists, that it is Lipschitz-continuous on the image $g([0, 1])$ and apply the mean value theorem on g^{-1}. The convergence rate follows then by a Taylor expansion.

In practice one may also be interested in a lower bound for the error. In the case (2.4.2), that is when $h(\pi, z)$ is positive, it is relatively simple to find a lower bound, which again turns out to be proportional to the variance of the small jumps. The Corollary 2.4.2 yields then that the error is bounded by $\sigma^2(\epsilon)$ from below and above, if (2.4.2) holds.

Going back to the case of an NIG Lévy process $L(t)$ and using (1.2.6), then neglecting the small jumps and solving the portfolio optimization problem yields in an error which can be bounded linearly in ϵ:

$$|\pi^*_{N,\epsilon} - \pi^*| \le C \times \epsilon \,.$$

A numerical example illustrates this in Figure 2.2. Using relevant parameters for the NIG process, we investigate the difference between the optimal strategy π^* and its approximation

$\pi^*_{N,\epsilon}$ for different $\epsilon > 0$. The truncated strategies are lower than the optimal ones meaning that the investor places less into the risky asset, although the truncation of the small jumps leads to less volatility in the driving noise. Here, less noise might be outweighted by a reduced return in the stock.
In Proposition 2.4.3 we show that the approximation using a Brownian motion leads to a convergence of $\pi^*_{W,\epsilon}$ to π^* with the same rate as for the case where small jumps are neglected. After this article was accepted for publication, we were able to improve the rate compared to the case when the small jumps are neglected. Therefore Section 2.4.3 is new compared to the accepted version. Here we find that

$$|\pi^*_{W,\epsilon} - \pi^*| \leq C_W \; \epsilon \, \sigma^2(\epsilon) \,,$$

for a constant $C_W > 0$ independent of ϵ. In fact, in the integral equation (1.2.8) the additional terms appearing from the Brownian approximation coincide with the second order terms of the Taylor expansion of h around $z = 0$ and they cancel each other in the proof. Choosing a NIG Lévy process for L leads then to a quadratic speed of convergence in ϵ

$$|\pi^*_{W,\epsilon} - \pi^*| \leq C \times \epsilon^2 \,.$$

We next investigate the convergence of the corresponding value functions. We first discuss maximization of terminal wealth, and then move on to analyse the convergence of the optimal consumption and the value functions. As we note in Section 2.2, maximising wealth over optimal investment and consumption pairs (π, c) results in the same optimal strategy π^* as when maximising expected utility over terminal wealth. As our results show, approximations of these control problems leads to convergence of the optimal investment strategies, as well as the value functions for maximization of the utility of terminal wealth. We show next that including consumption does not alter these conclusions. Still remaining is the convergence of the wealth processes. Convergence in probability of the wealth processes is clear. To prove convergence in $\mathcal{L}^2$ we use Cauchy-Schwarz and Gronwall's inequality and can derive a rate which is again proportional to $\sigma^2(\epsilon)$.

The convergence and convergence rates in this paper are analysed for the specific case of power utility in a Merton framework. The proofs, especially for the convergence rate of the optimal control, depend on features of the concrete form of the solutions in this specific setting. In a more general setting a concrete solution is not available. Additionally it is not clear if the optimal control and the consumption rate are constant in time.

1.2.2 Summary of Article 2: "On Stability to Model Risk of Options in a Bivariate Lévy Market"

We consider a bivariate Lévy process $L = (L^{(1)}, L^{(2)})$ with Lévy measure $\nu(dz) = \nu(dz_1, dz_2)$, where the components have Lévy-Itô decomposition

$$\begin{aligned} L^{(1)}(t) &= \textstyle\int_0^t \int_{|z|<1} z_1 \tilde{N}(ds, dz_1, dz_2) + \int_0^t \int_{|z|\geq 1} z_1 N(ds, dz_1, dz_2) \\ L^{(2)}(t) &= \textstyle\int_0^t \int_{|z|<1} z_2 \tilde{N}(ds, dz_1, dz_2) + \int_0^t \int_{|z|\geq 1} z_2 N(ds, dz_1, dz_2) \,, \end{aligned} \tag{1.2.9}$$

with $|z| = \sqrt{z_1^2 + z_2^2}$. Analogous to the one dimensional case, define

$$\sigma_1^2(\epsilon) = \int_{|z|<\epsilon} z_1^2 \nu(dz_1, dz_2) \tag{1.2.10}$$

$$
\begin{aligned}
\sigma_2^2(\epsilon) &= \int_{|z|<\epsilon} z_2^2 \nu(dz_1, dz_2) \\
\sigma_{12}(\epsilon) &= \int_{|z|<\epsilon} |z_1 z_2| \nu(dz_1, dz_2) \ ,
\end{aligned}
$$

and set

$$
\Sigma(\epsilon) = \begin{pmatrix} \sigma_1^2(\epsilon) & \sigma_{12}(\epsilon) \\ \sigma_{12}(\epsilon) & \sigma_2^2(\epsilon) \end{pmatrix} \ .
$$

We approximate the jumps with size $|z| < \epsilon$ by a bivariate scaled Brownian motion, where the scaling matrix

$$
\alpha(\epsilon) = \begin{pmatrix} \alpha_1(\epsilon) & \alpha_2(\epsilon) \\ \alpha_2(\epsilon) & \alpha_3(\epsilon) \end{pmatrix}
$$

is chosen such that the approximated process has the same variance as the original one. This is equivalent to

$$
\alpha(\epsilon)\alpha(\epsilon)^T = \Sigma(\epsilon) \ .
$$

Then we have $L_\epsilon = (L_\epsilon^{(1)}, L_\epsilon^{(2)})$ as an approximation of $L = (L^{(1)}, L^{(2)})$ with

$$
\begin{aligned}
L_\epsilon^{(1)}(t) &= \alpha_1(\epsilon)B^{(1)}(t) + \alpha_2(\epsilon)B^{(2)}(t) \\
&\quad + \int_0^t \int_{\epsilon \le |z| < 1} z_1 \tilde{N}(ds, dz_1, dz_2) + \int_0^t \int_{|z| \ge 1} z_1 N(ds, dz_1, dz_2) \\
L_\epsilon^{(2)}(t) &= \alpha_2(\epsilon)B^{(1)}(t) + \alpha_3(\epsilon)B^{(2)}(t) \\
&\quad + \int_0^t \int_{\epsilon \le |z| < 1} z_2 \tilde{N}(ds, dz_1, dz_2) + \int_0^t \int_{|z| \ge 1} z_2 N(ds, dz_1, dz_2).
\end{aligned} \tag{1.2.11}
$$

First, we consider option prices of the form

$$
C_X = \mathbb{E}[f(X(T))] \ ,
$$

where X is a jump-diffusion in $\mathbb{R}^2$ and $f : \mathbb{R}^2 \to \mathbb{R}$ a payoff function. Eberlein, Glau and Papapantoleon [34] derive pricing formulas based on the Fourier transform $\widehat{f}$ of the payoff function f. They write the option price as

$$
C_X = \frac{e^{-R.s}}{(2\pi)^2} \int_{\mathbb{R}} e^{-iu.s} \phi_{X_T}(u - iR) \widehat{f}(iR - u) du \ , \tag{1.2.12}
$$

with $s \in \mathbb{R}^2$, a damping factor $R \in \mathbb{R}^2$ and where $R.s$ denotes the scalar product between R and s. Using (1.2.12), the difference between the option prices written on L and L_ϵ is given by

$$
|C_L - C_{L_\epsilon}| = \left| \frac{e^{-R.s}}{(2\pi)^d} \int_{\mathbb{R}} e^{-iu.s} \Big(\phi_{L(T)}(u - iR) - \phi_{L_\epsilon(T)}(u - iR) \Big) \widehat{f}(iR - u) du \right| \ .
$$

Therefore, we concentrate on estimating the difference between the characteristic functions $\phi_{L(T)}$ and $\phi_{L_\epsilon(T)}$. In the Taylor expansions of the Lévy integral of the characteristic exponent of $L(T)$, the second order terms coincide with the Brownian term of the characteristic exponent of $L_\epsilon(T)$. This leads to a convergence rate proportional to

$$|C_L - C_{L_\epsilon}| \leq d_1 \epsilon \sigma_1^2(\epsilon) + d_2 \epsilon \sigma_2^2(\epsilon) + d_{12} \epsilon \sigma_{12}(\epsilon) \; , \tag{1.2.13}$$

for constants d_1, d_2 and d_{12}. Here $\sigma_{12}(\epsilon)$ corresponds to the covariance and is the result of including dependency. If one only truncates the jumps without approximating them, we get a convergence rate proportional to

$$|C_L - C_{L_{N,\epsilon}}| \leq d_1 \sigma_1^2(\epsilon) + d_2 \sigma_2^2(\epsilon) + d_{12} \sigma_{12}(\epsilon) \; ,$$

In Section 3.3 we consider spread options and state a version of Margrabe's formula when the log returns of the underlying are driven by (1.2.9). We start under the real world measure $\mathbb{P}$, define by Esscher transforms equivalent measures, and fix a parameter θ such that $\mathbb{Q}_\theta$ is risk neutral. Also Margrabe's pricing measure $\mathbb{Q}_{\theta+\mathbf{1}_1}$ is derived with an Esscher transform. If one approximates under $\mathbb{Q}_{\theta+\mathbf{1}_1}$, we get a convergence rate of the form (1.2.13). If one approximates under the real world measure $\mathbb{P}$, then the approximated asset price dynamics depend on ϵ. Therefore, the risk neutral measure that turns the discounted approximated asset prices into martingales dependents also on the truncation level ϵ. Denote the corresponding Esscher parameter by θ_ϵ. This setting was studied by Benth, Di Nunno and Khedher [10], and they find a convergence rate of the Esscher parameters proportional to $\sigma^2(\epsilon)$, the variance of the small jumps. They find the same rate also for the option prices. Here, we use their result on the Esscher parameter and find a convergence rate as a sum of several terms. Our proof allows to see what parts of the convergence rate are due to the change of measure and what is due to the Brownian approximation. After a measure change, the approximation (1.2.11) is not 'perfect' anymore, in the sense that the variance of the small jumps does not stay the same under the new measure. Then, the second order terms in the characteristic exponents do not coincide anymore. Nevertheless, it is possible to get an estimation of their difference. In the convergence rate, we consider the variation in terms of the new measure. As the proofs in Article [2] are different from the ones in Benth, Di Nunno and Khedher [9, 10, 11], we can see the effect of adding the Brownian approximation in the quantification of the difference of the option prices, compared to the case where one simply truncates the small jumps.

Article [2] is a completely revised version of Benth et al. [12]. There, we prove Margrabe's formula for exponential jump diffusions with stochastic factors. For these dynamics we prove convergence of Margrabe's formula when approximating the small jumps, but we could not determine a rate of convergence. Therefore, Article [2] goes back to the case where we were able to specify convergence rates and see the effect of including dependency into the model.

1.2.3 Conclusion: Brownian approximation

For the convergence of the optimal control in Article [1], the option prices in Article [2] and the $\mathcal{L}^1$ convergence of $L_\epsilon = (L_\epsilon^{(1)}, L_\epsilon^{(2)})$ in Proposition 3.2.1, it is only necessary that the scaling coefficients of the Brownian motion converge to 0. This admits a much wider range of possible scaling coefficients then those in (1.2.4) and (1.2.10), that keep the variance of the original process. The convergence results hold even if one does not approximate at all and only truncates the small jumps. Another issue is the resulting quality of the approximated optimal control and option prices. An approximation of the small jumps in the underlying driving Lévy process should lead to a significant improvement of the approximated option price, compared to the case where one only truncates the small jumps. For the error estimate, it is important how the scaling coefficients look like.

We find the same type of convergence rate both for the optimal control and option prices.

1.3 Energy markets

A world-wide liberalization of electricity markets started in the beginning of the 1990s. In Norway the parliament decided to deregulate the market for power in 1991. This resulted in the establishment of what is known today as NordPool, one of the first energy exchanges. In 2002 it was established what became the EEX, the European Energy Exchange. The liberalization has given rise to a number of new markets with distinctive features that need special attention when models are set up. For instance, one observes big spikes in the spot prices, that are due to imbalances in supply and demand. In practice these spikes lead to big price risk for electricity producing and consuming companies, so that new tools and instruments where needed for risk management. Futures and forwards on the spot are traded frequently, and options on the futures are available as well.

For example, the electricity market of the EEX can be divided into three submarkets: a day-ahead spot market, a financial market for futures contracts on power, and a market for plain vanilla call and put options on the futures. The power spot prices are determined in an auction-based system, where the traders hand in prices and volumes for production or consumption for given hours the next day. Based on these bids, the exchange creates demand and supply curves for each hour the following day, and at 2 pm the EEX publishes these spot prices for the next 24 hours of the next day. One distinguishes between base and peak load prices. The base load prices are settled as average over all 24 hours at all days of the week, while the peak load prices take only the peak hours into account, which are from 8 in the morning until 8 in the evening at working days.

The future contracts deliver the power over an agreed period of time for an agreed price. Therefore they function as a swap contract, where each day during the delivery period the floating spot price is exchanged against the fixed forward price. There are markets, where both forwards and futures are traded. Here, we shall not make a distinction between these two asset classes.

The EEX offers furthermore European style put and call options written on a specified futures contract, where the exercise takes place four trading days prior to the beginning of the delivery period of the underlying futures. They are written on Phelix base load futures with monthly, quarterly and yearly delivery periods.

On their webpage, the EEX gives descriptions and examples on how the considered contracts can be used by market participants. The spot market is used by trading participants to optimize the purchase and sale of quantities of power in the short-term, that is to trade electricity with delivery on the same or next day. They can use the products on the futures market in order to hedge against the risks of future price changes in the long run. For this purpose there are also European call and put options available. For instance, consider a producing

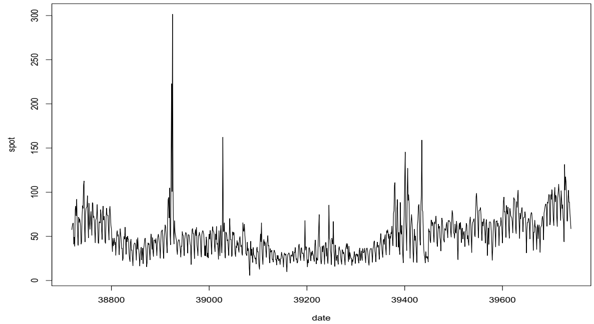

Figure 1.1: EEX Phelix base load spot prices from 02.01.2006 - 19.10.2008.

company with a high use of energy. If they purchase the electricity at the spot market with its big price variations and extreme spikes, a not negligible risk factor is connected to energy costs. To close the risk position connected to the purchase of energy, they can buy the needed energy for the following month for a fixed price on the futures market. Assume furthermore that the company only produces efficiently, if the energy cost is below a certain level. Here, options can be applied. By buying a call option the company secures a maximum price to buy electricity corresponding to the exercise price of the option (plus the option premium). Although not traded at the EEX, we remark that spread options considered in Article [2] naturally appear in Energy markets. Consider for example a power plant that uses gas in order to produce electricity. It purchases gas and sells electricity and is therefore interested in the price spread between the two products.

The electricity markets challenge academic research, as it turns out to be difficult to propose well-fitting models for the spot. Recent streams of literature describe the spot price with CAMRA processes (see Garcia et al. [39], Benth et al. [16]), or Lévy semistationary processes (see Barndorff-Nielsen et al. [7]). For a comprehensive text book on stochastic modeling of electricity markets, see Benth et al. [19]. The models for the spot price that we use in Article [3] and Article [4] are based on the two-factor model for commodity prices from Schwartz and Smith [55] (see also Gibson and Schwartz [42]). It was applied to electricity spot prices by Lucia and Schwartz [47]. In Article [3] and [4] we do not concentrate on improving the pathwise fit of spot and futures, but are interested in option pricing. Therefore we choose tractable, established models, that might not be as sophisticated as the ones that where recently suggested, but provide a satisfying distributional fit. After specifying a model for the spot, one typically defines the futures price as derivative from the spot. Another approach is to model the futures curve directly following the Heath-Jarrow-Merton approach from interest theory (see Benth et al. [19], Benth and Koekkebakker [17]). For an overview over forward curves in commodity markets, see Borovkova [24]. While there is plenty of literature that studies topics around spot and futures, the literature about pricing and hedging of energy options is not so extensive. For the case of no jumps see Benth et al. [19].

A stylised feature of empirical spot prices is the occurrence of spikes and big price variations, see for example Figure 1.1 for EEX base load spot prices from 2006 until 2008. They are caused by unexpected imbalances in supply and demand of energy. For example, an un-

predicted drop of the temperature leads to more heating and a sudden increased demand in energy, and results in a positive spike. On the other hand, negative spikes are observed as well in connection with sudden overproduction of electricity. This can appear in connection with renewable energy sources as wind energy. An unexpected sudden increase of wind results in an unwanted overproduction of electricity.
The two-factor model after Schwartz and Smith [55] captures the spiky behaviour in a separate component, which is modeled by a mean reverting Ornstein-Uhlenbeck process. Typically, it takes only some days until the spot price is back on its original level, so that the rate of mean reversion for the big spikes is very high. One is now interested in the effect of the spike component in the derivatives of the spot. This is analysed for the implied forward dynamics in Benth et al. [19]. In Article [3], we go further and quantify the impact of the spikes on the option prices.

Another characteristic of electricity is that it is not storable. The trading in the spot market is physical, and after producing the energy it needs to be consumed immediately. There are exceptions as producers may for example store energy in water reservoirs, but this applies only to a part of the market, and is not possible on the consumers side either. Thus, the electricity spot is not a tradable asset in the sense that one cannot form financial portfolios with it. As an implication, it is not possible to apply the no-arbitrage theory to price derivatives based on the spot. Therefore the measure to price the future on the spot does not need to be a martingale measure. In Article [4] we discuss direct implications of the break down of the no-arbitrage theory when it comes to option pricing.

1.3.1 Summary of Article 3: "Pricing and Hedging Options in Energy Markets by Black-76"

Many models for the spot price evolution divide the price evolutions into a short-term and a long-term component and take mean reversion of the electricity price as well as uncertainty in the equilibrium level to which the prices revert into account. The popular Schwartz and Smith model for commodities (see Schwartz and Smith [55]) is such a model. The non-stationary long-term factor models the equilibrium price level, and reflects expectations of improving technologies for the production of the commodity, inflation or political and regulatory effects, and depletion of non-renewable resources like gas and coal. The mean reverting short-term factor describes changes in demand and supply resulting for example from variations in the wether conditions and sudden outages of power plants. Electricity markets are well-known for their large price variations and rare, big spikes, which are captured by the short-term component. In this article, we examine the influence of the short-term and long-term factor on the spot and prove that the short-term factor is insignificant for pricing options in many relevant cases. Furthermore, we determine the error between the option prices with and without short-term factor as exponentially depending on the speed of mean reversion of the short-term factor and the time between exercising the option and delivery of the forward.

Typically forwards deliver the underlying energy over a delivery period $[T_1, T_2]$, say a month or a year. Then the forward price should be defined as expected average spot over the delivery period

$$F(t, T_1, T_2) = \mathbb{E}_Q \left[\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} S(t)\, dt \,|\, \mathcal{F}_t \right] . \tag{1.3.1}$$

The delivery period smoothens the effect of the spikes in the forward price and should thus also lower their influence in the option price. When the spot model is exponential, there is no analytical closed form solution for the forward price (1.3.1), which would make the examination of the price of an option with the forward as underlying barely tractable. One alternative is to choose some point in the delivery period, for example the mid-point $T = (T_1 + T_2)/2$, and to approximate the delivery period by this point of time. Then the standard definition of the forward price $f(t,T)$ at time $t \geq 0$ of a contract delivering the underlying energy at time $T \geq t$ is

$$f(t,T) = \mathbb{E}_Q[S(T) \,|\, \mathcal{F}_t] \,,$$

for some pricing measure Q being equivalent to the real world measure P.
Often, the exercise time τ of an option is shortly before the beginning of the delivery period T_1. By approximating the delivery period by its mid-point, one constructs formally a time $(T-\tau)$ from exercising the option until the delivery of the forward of at least half the delivery period. The time $(T-t)$ from purchasing the forward until its (constructed) delivery is even longer. For monthly or yearly periods, this can be long enough to be "long time to delivery", where the influence of the short-term factor does not matter any more. The approximation of the delivery period incorporates the well-known empirical fact that in electricity markets forward prices do not converge to the spot prices if one approaches the delivery time.

In this paper, the energy spot price follows an exponential two-factor model of the form

$$S(t) = \Lambda(t) \exp(X(t) + Y(t)) \,,$$

where the non-stationary long-term factor X is a drifted Brownian motion

$$dX(t) = \mu \, dt + \sigma \, dB(t) \,.$$

The stationary short-term factor Y is an Ornstein-Uhlenbeck process with dynamics

$$dY(t) = -\beta Y(t) \, dt + dL(t) \,.$$

It is driven by a pure jump Lévy process L and the parameter β describes the speed of mean reversion. β plays a central role as it will determine the speed of convergence. We shall assume that L and B are independent to simplify the analysis. In the Gibson and Schwartz model [42], both factors have Gaussian stochastic drivers. Considering the large price deviations in the electricity spot prices this assumption seems to be unrealistic here, especially for the driver of the short time factor. We choose to use a pure jump Lévy process instead. We need to assume that the Lévy process has finite exponential moments until a certain order, and it turns out that we need to require finite moments up to order three for the analysis to be well defined.
We assume that the dynamics are already stated under a pricing measure Q. In Proposition 4.2.1 we derive an explicit expression of the forward price at time t in terms of $X(t)$, $Y(t)$ and the logarithmic moment generating function of $L(1)$. Using Itô's formula for jump processes, the forward price dynamics turns out to be a geometric jump diffusion (see Proposition 4.2.2) with dynamics

$$\frac{df(t,T)}{f(t-,T)} = \sigma \, dB(t) + \int_{\mathbb{R}} \left\{ \exp\left(z e^{-\beta(T-t)} \right) - 1 \right\} \widetilde{N}(dz, dt) \,. \tag{1.3.2}$$

Without the jump component, the forward dynamics (1.3.2) turn out to be a geometric Brownian motion, the underlying dynamics of the forward price in the Black-76 model. We shall compare the option price with underlying two-factor model for the spot with the Black-76 price, where only the long-term factor is taken into account.

We consider European call options on the forward with the geometric forward price dynamics (1.3.2). In Proposition 4.3.2 we derive a Black-76-type representation of the option price by conditioning on the jump part and evaluating the inner expectation by the Black-76 formula. This representation is the basis to prove Theorem 4.3.8, the main result of the paper. It states that the option price converges uniformly to the Black-76 price when the time to delivery goes to infinity. That is, we have

$$\sup_{x\geq 0} |C(t,\tau,T,x) - C_{\mathrm{B76}}(t,\tau,T,x)| \leq c\mathrm{e}^{-\beta(T-\tau)}\,,$$

for some constant c and $\tau < T$. The convergence is exponential, where the speed of convergence is determined by the mean reversion parameter β of the short-term component and the time between exercising the option and the delivery of the forward. We prove this convergence result by a sequence of Lemmas, that derive the same convergence rate for smaller parts of the option price. Their proofs are technical and use the series representation of the exponential function in combination with the Cauchy-Schwartz inequality, the mean value theorem, the Esscher transformation and the dominated convergence theorem. The log-moment generating function of $L(1)$ helps to keep the notation short. Theorem 4.3.8 follows then using the triangle inequality.
A numerical example emphasizes the practical relevance of Theorem 4.3.8. For the pure jump Lévy process, we chose a compound Poisson process with exponential jump size distribution and parameters, that are realistic in the winter month in the Nordic electricity market NordPool. We calculate the call option price after Proposition 4.3.2 by Monte Carlo simulation and plot the difference to the Black-76 price for an option at the money (see Figure 4.1). It exercises in 10 days, and the plot shows the absolute value of the difference in dependence of the delivery point of the forward. We see exponential convergence of the numerical calculated prices and add the theoretical error bounds. In the scenario discussed above, namely for a monthly delivery period of 30 days which is approximated by its midpoint, the forward delivers $T-\tau = 15$ days after the exercise of the option, that is $T-t = 25$ days after entering the option. At this point of time the prices are already very close and we have a mispricing of only 3 %. We conclude that here the Black-76 formula is a good approximation.

Given the popularity of Fourier transform based pricing methods for options, we include a pricing formula based on Fourier transform for the two expectations in Proposition 4.3.2. As the expressions in the expectations are not in $\mathcal{L}^1(\mathbb{R})$, we dampen them using an exponential function, prove that they are dampened integrable on $\mathbb{R}$ and compute the Fourier transform of the damped functions.

The last chapter considers quadratic hedging of call options on forwards. As the forward price dynamics is a jump diffusion process, the market is incomplete and it is not possible to hedge the option completely by a portfolio consisting of the underlying forward contract and a bank account. We choose quadratic hedging instead, that minimises the variance of the hedging error, defined as the difference between the terminal value of the hedging portfolio and the option. Compare Cont and Tankov [31] for an overview over hedging in incomplete markets. First, we derive the quadratic hedging position an time $t \leq \tau$ using Itô calculus and the Itô isometry for jump processes. Finally we show in a sequence of lemmas, that the quadratic hedging position converges uniformly to the delta hedge, with an exponential rate given by β. That is, we have the same speed of convergence as we derived for the option prices.

1.3.2 Summary of Article 4: “Pricing Futures and Options in Electricity Markets“

Given two derivatives with tradable underlying assets in a market, the no-arbitrage theory requires that there needs to exist a pricing measure Q such that both assets are discounted martingales with respect to this measure. Otherwise, arbitrage is possible. The standard approach is then to price both derivatives under Q. However, it is possible to price the derivatives with two different measures, such that the underlying assets are discounted martingales with respect to the corresponding pricing measure. Nevertheless, the two pricing measures still have to fulfil the overall no-arbitrage condition that will connect them and tie them together. If one of the underlying assets is not tradable in the financial sense, the pricing measure for this asset does not need to be a martingale measure, the no-arbitrage conditions break down and the pricing measures are no longer tied together. They can be completely disconnected from each other. This is the case in our setting. Electricity is not storable, so that it must be consumed once it is purchased. It is not possible to form a financial portfolio, that is based on buying the asset and selling it again at a later point of time. In this sense, the electricity spot is not a tradable asset and the no-arbitrage argumentation that is based on cost and carry strategies (see Duffie [33]) cannot be applied. Therefore one can use different measures to price the futures and the option. The goal of this paper is to examine empirically wether or not the traded options are priced under the same pricing measure as the futures. Before doing so, we need to specify a model for spot and futures and fit it to data.

We choose an arithmetic model for the spot price in the spirit of Lucia and Schwartz [47], as the occurrence of negative prices for the spot in electricity markets suggests. The electricity spot price at time t is then given by

$$S(t) = \Lambda(t) + X(t) + Y(t)\,.$$

Here, $\Lambda : [0, \widetilde{T}] \mapsto \mathbb{R}$ is a measurable deterministic function, modeling the mean seasonal variation in spot prices. The base (or long-term) component $X(t)$ in the spot price dynamics is assumed to be non-stationary and defined to be a Lévy process

$$dX(t) = dL_1(t)\,.$$

Furthermore, $Y(t)$ is a mean- reverting short-term factor of the form

$$dY(t) = -\eta Y(t)\,dt + dL_2(t)\,.$$

Here, the constant $\eta > 0$ is expected to be rather big, since the spikes created by the Lévy process $L_2(t)$ are reverting fast back to the normal price level. After specifying the spot model, we introduce a parametric class Q_θ of equivalent probability measures via Esscher transformation and determine the Lévy process under the measure Q_θ. Proposition 5.3.2 gives the implied futures price under Q_θ as

$$F(t, T_1, T_2) = \bar{\Lambda}(T_1, T_2) + X(t) + Y(t)\bar{\eta}(t, T_1, T_2)$$

$$
\begin{aligned}
&-\frac{1}{2}\mathrm{i}\psi_x(-\mathrm{i}\theta)(T_2-T_1)-\mathrm{i}\psi_x(-\mathrm{i}\theta)(T_1-t) \\
&+\frac{-\mathrm{i}\psi_y(-\mathrm{i}\theta)}{\eta}(1-\overline{\eta}(t,T_1,T_2))\,,
\end{aligned}
$$

where $\overline{\eta}(t,T_1,T_2)$ is defined in the proposition and ψ_x and ψ_y denote the partial derivatives of the log-characteristic function ψ with respect to the two variables x and y. Proposition 5.3.3 states the corresponding dynamics.

Section 5.4 describes the empirical study of the EEX spot and futures data. Our method estimates the parameters in the spot model as well as the market price of risk in the same procedure. The long-term component $X(t)$ and the short-term component $Y(t)$ are not directly observable. Nevertheless, they can be estimated with the help of futures data using the result in Proposition 5.4.1. It states that asymptotically futures prices behave like

$$
F(t,T_1,T_2)\approx\overline{\Lambda}(T_1,T_2)+\Psi(t,T_1,T_2;\theta)+X(t)\,, \tag{1.3.3}
$$

for $T_1-t\to\infty$. Here, $\Psi(t,T_1,T_2;\theta)$ is deterministic and defined in the proposition and $\bar{\Lambda}(T_1,T_2)$ is the average value of the seasonality function $\Lambda(s)$ over the interval $[T_1,T_2]$. This means that in the long end of the forward market the forward price fluctuates as the non-stationary factor $X(t)$ plus some non-stochastic adjustment term $\bar{\Lambda}(T_1,T_2)+\Psi(t,T_1,T_2;\theta)$ involving the market price of risk θ. The short-term factor $Y(t)$ does not contribute significantly to the futures price for long time to delivery, so that $X(t)$ can be filtered from the forward prices using (1.3.3). This approach is analogous to the calibration procedure in Schwartz and Smith [55] and a more sophisticated version can be found in Benth et al. [16]. We worked with a small data set from the EEX and had available daily Phelix base load spot prices from 02.01.2006 until 19.10.2008 as well as base load future contracts with 1 month delivery. After deseasonalising the spot, we create a time series of futures price data with long time to delivery. We derived $\widehat{T}=16$ as the threshold when $Y(t)\overline{\eta}(t,T_1,T_2)\approx 1$ using $Y(t)$ being three times the standard deviation of spot price data. This depends obviously on the value of η, which is the speed of mean reversion of the process Y. We can estimate this parameter from the autocorrelation function of Y, which is known to be exponentially decaying at the rate η (see Benth et al. [19]). However, at this point in the estimation procedure we had not yet filtered the time series of Y from the spot data, so the empirical autocorrelation function was unknown to us. Therefore, we did a rough estimation of η by looking at the empirical autocorrelation of the deseasonalised spot, which is modeled by $X(t)+Y(t)$. We observed a decaying autocorrelation structure, and fitted an exponentially decaying function to the first five lags obtaining a pre-estimate $\widehat{\eta}$. Next, we estimated the deterministic adjusting terms in (1.3.3) by linear regression and filtered $X(t)$ from the deseasonalised forward prices. Then we could fit a Lévy process to the residuals of the Y process and the time series of the X process obtained above. From the fitted Lévy process we obtained the cumulant Φ from which we found the market price of risk as solution of the equations in (5.4.2).

The autocorrelation function of the filtered time series $Y(t)$ of the short-term component shows exponential decay and is shown in Figure 5.2. For an Ornstein-Uhlenbeck process the theoretical autocorrelation function is exponential decaying at rate η. A single exponential function is able to capture either the beginning of the theoretical autocorrelation function, or the end. This indicates that it might be a better choice to use two Ornstein-Uhlenbeck processes, one for the extreme spikes with extremely quickly mean reversion and one for the more normal spikes, which mean revert slower. For simplicity of the study, we stick to only

one Ornstein-Uhlenbeck process. Like that we won't replicate the spiky behaviour exactly, but as we are interested in option pricing, we focus more on distributional than pathwise properties.
The filtered component of the long time factor appears to be clearly non-Gaussian, although it is often assumed to be a driven by a Brownian motion. We find the fit of the normal inverse Gaussian (NIG) process satisfying for our purposes, which we choose for both driving processes L_1 and L_2. The empirical and estimated densities are shown in Figure 5.3. Our estimation results in a positive market price of risk $\theta = (\theta_1, \theta_2)$, which shifts the NIG distributions towards the right. Indeed, while the expected values of $L_1(1)$ and $L_2(1)$ are negative under P, they are positive under Q_θ. For a summary of the estimated parameters of the NIG distributions and the market price of risk, see Table 5.2 and Table 5.3.

The market for options is very illiquid at the EEX. Table 5.4 and 5.5 give the main characteristics of the four call options and seven put options that where traded from January until October 2008. At the EEX, options exercise four trading days before the delivery period of the underlying futures starts. Options are traded only with underlying base load future contracts, which explains why we chose to work with base load data. In fact, the illiquidity of the considered option market is an issue that can question our analysis.
The classical approach in commodity markets is to use the Black-76 formula to price options on futures. The famous formula is stated in Proposition 5.5.1 and assumes that the futures price dynamics is given by a geometric Brownian motion. Using the historical volatility of the last month of the underlying futures contract, the calculated Black-76 prices underestimate substantially the quoted option prices at the EEX. See Table 5.6 and 5.7 for an overview of the results. A first conclusion might be, that the Black-76 formula is not suitable for option pricing in its basic form, if one uses the volatility of the underlying futures as input. One might speculate that the quoted prices include a big risk premium for effects like illiquidity and an adjustment for non-Gaussianity in the futures price dynamics.
As our proposed futures price dynamics is far more sophisticated than a simple geometric Brownian motion, we move on to analyse the implied option prices from our model. The call option price under the measure Q_θ is given by

$$C(t) = e^{-r(\tau - t)} \mathbb{E}_{Q_\theta}[F(\tau, T_1, T_2) - K \mid \mathcal{F}_t]\,. \tag{1.3.4}$$

The Q_θ-dynamics of $F(t, T_1, T_2)$ is given by Proposition 5.3.3, where Q_θ is determined through the market price of risk (5.4.4) from the empirical spot-futures study. We evaluate the expectation in (1.3.4) through a Monte-Carlo simulation based on 1,000,000 paths. To simulate the NIG distribution, we applied the algorithm implemented in the R-package fBasics, which is based on the normal variance-mean mixture of the NIG distribution. The results are reported in Table 5.8 and Table 5.9.

In the discussion of the results, we take the time to maturity of the options into account. In the theoretical futures dynamics the spike component does not have much influence for long time to maturity, and one is left with the base component. However, the base component is filtered directly from the futures data with long time to maturity. If our futures model is correct and the options are priced using the inherited risk perception from the futures, then options with long time to maturity should tend to be priced accurate and errors that might result from a misspecification of the spike component should be excluded as much as possible.
Another point for the analysis is whether the options are in, at or out of the money, as it

influences the proportion of the probability distribution that is taken into account. If the distribution used for pricing (that is, the pricing measure) is not correct, this should lead to a bigger misspecification for those options, where a lot of the distribution is taken into account. For options at the money, a big part of the distribution is included, so that the misspecification due to a wrong pricing measure should be biggest.
The call options C1 and C2 (see Table 5.4) have the longest time to maturity and are about at the money. If the market prices according to our forward model and uses the probability Q_θ, they should be priced reasonably accurate. However, the simulated option prices using our suggested model are about 50% higher than the quoted ones. This means that our model is pricing in too much risk. From the spot dynamics we estimated positive market prices of risk which pushes the skewness of the NIG distribution to more positive jumps. The more positive the market price of risk is, the higher are the values of the options. Thus, it seems like the forward model inherits far too much risk premium from the spot when it comes to option pricing. We reach the conclusion that the option market is not including the same risk perception as the one inherited from the spot in the futures market. The other simulated option prices support the conclusion (see Table 5.8 and Table 5.9). This would mean that a completely different pricing measure is used in the option market than in the futures market. We leave the suggestion of a suitable choice of measure for future research.

1.3.3 Discussion: Options on Futures in Energy Markets

In both articles [3] and [4] we use a two-factor model for the spot, which divides the spot price evolution into a short-term and a long-term component. Nevertheless, in Article [3] we use a geometric model and Article [4] is based on an arithmetic model. This takes the advantages and disadvantages of both models into account and uses the model which provides the best benefit for the given purposes.
The arithmetic model is able to capture negative prices, which are observed in the electricity spot markets. That makes it suitable for empirical studies, and we choose it for Article [4] that explores option prices based on data from the EEX. Furthermore, the arithmetic model leads to an analytical closed form solution for the futures price if one includes the delivery period in the futures price dynamics. This is not the case for the geometric model.
On the other hand, the Black-76 formula is based on a geometric Brownian motion for the futures dynamics. If one starts with an exponential two-factor model for the spot price, the dynamics for the forward price becomes a geometric jump diffusion. Ignoring the jump term results in our setting in a geometric Brownian motion, and the claim that for long time to delivery, when the influence of the short-term component is small, one can do well using the Black-76 formula. Here we approximate the delivery period by a point inside the period, as for example the mid-point.

The Black-76 formula appears in both papers. While [4] concludes that it is not the classical form of the Black-76 formula that is used in the market, [3] suggests that the Black 76 formula gives a good approximation for the option price if one uses our suggested geometric two-factor model for the spot. This seems to be contradictory.
In Article [3] the quality of the approximation of the option price depends on two parameters, the time to delivery and the speed of mean reversion of the short-term component. Here, a fast mean reversion in the underlying spot is crucial. The autocorrelation function in [4] indicates that one might improve the fit of the spot model by adding a second Ornstein-

Uhlenbeck component with slower mean reversion. The question is how this influences the convergence rate in Article [3]. It is possible that one would need to increase the time to delivery to compensate a slower rate of mean reversion. In the example in Section 4.3 we consider options on futures with a monthly delivery period which we approximate by its midpoint. Then we have 15 days until delivery and a fast speed of mean reversion. In this scenario, the Black-76 formula is a good approximation. A slower rate of mean reversion might not lead to a good approximation for monthly delivery periods.
The difference in the spot models makes a direct comparison difficult. However, in [3] we assume that the long-term component is Gaussian, while the filtered arithmetic long-term component [4] is far away from being Gaussian. One could suspect that a filtered corresponding geometric long-term component might not be Gaussian either.

In Article [3] we consider the impact of big price variations, while in the Article [1] and [2] we look at the behaviour of small variations. There, the error estimation depends again on the behaviour of the small variations and is characterized in terms of the Lévy measure. In Article [3] the characterisation of the behaviour of the big price variations of the driving noise (as for example the behaviour of the tail of the Lévy measure) does not influence the rate of convergence. Nevertheless, the size of the jumps will influence the constant c in Theorem 4.3.8 and so influence the error.

CHAPTER 2

ARTICLE 1: "STABILITY OF MERTON'S PORTFOLIO OPTIMIZATION PROBLEM FOR LÉVY MODELS"

Fred Espen Benth and Maren Diane Schmeck

Abstract

Merton's classical portfolio optimisation problem for an investor, who can trade in a risk-free bond and a stock, can be extended to the case where the driving noise of the log-returns is a pure jump process instead of a Brownian motion. Benth et al. [13], [14] solved the problem and found in the HARA-utility case the optimal control implicitly given by an integral equation. There are several ways to approximate a Levy process with infinite activity: by neglecting the small jumps or approximating them with a Brownian motion, as discussed in Asmussen and Rosinski [5]. In this setting, we study stability of the corresponding optimal investment problems. The optimal controls are solutions of integral equations, for which we study convergence. We are able to characterize the rate of convergence in terms of the variance of the small jumps. Additionally, we prove convergence of the corresponding wealth processes and indirect utilities (value functions).[1]

2.1 Introduction

In Merton's [49] seminal paper on optimal portfolio management under uncertainty, it is proved that a risk averse investor will place a constant proportion of her total wealth in risky assets. Optimality is measured as the expected utility of terminal wealth, with a power or HARA utility function measuring the risk preferences of the investor. Moreover, the dynamics of the risky assets follow a geometric Brownian motion. The optimal proportion is

[1]To appear in: Stochastics. Section 2.4.3 is completely new in comparison with the accepted version.

given explicitly as the ratio of the excess return over the risk free, normalized by the volatility of the risky asset and the risk aversion of the investor.

The constant proportion rule is among the popular strategies for portfolio management in practice. Merton's portfolio selection problem has also gained a lot of attention in the scientific literature over the years, with generalizations in various directions. Recent extensions of the original Merton problem include the case of stochastic coefficients in Delong and Klüppelberg [32], and bounded downside risk through restrictions on Value-at-Risk and Expected Shortfall in Klüppelberg and Pergamenchtchikov [45]. For a general treatment and discussion, we refer to Øksendal and Sulem [58].

One stream in the literature relevant for our considerations focuses on analysing the effects of more realistic models for the risky asset price dynamics on the optimal portfolio management problem. For example, Benth, Karlsen and Reikvam [13] examined Merton's problem when the risky asset price dynamics is given by an exponential pure-jump Levy-process. Also in this case it is optimal to invest a constant proportion of the investor's wealth in the risky asset, however, the proportion is given in terms of a solution of an integral equation involving the excess return of the asset and the characteristics of the jump processes. One may easily include consumption into the portfolio problem, regaining qualitatively similar solutions as in the classical Merton's problem. Emmer and Klüppelberg [37] examine constraints of an upper bound for the risk when stock prices follow an exponential Lévy process. There is also work on optimal portfolios with HARA-utility in multidimensional cases, see for example Callegaro and Vargiolu [27], where assets are driven by a multidimensional Poisson process, or Pasin and Vargiolu [57] in the case of exponential additive processes.

Lévy processes are popular in financial modelling since they are able to explain many of the stylized facts of asset prices (see Cont and Tankov [31] for a discussion of Lévy processes in finance). In particular, some processes like the normal inverse Gaussian (NIG) or the hyperbolic Lévy process have become particularly relevant since they are able to capture the return distribution of most asset prices (see e.g. Barndorff-Nielsen [6] and Eberlein and Keller [35]). These Lévy processes are pure-jump, and therefore give distinctively different paths of the asset prices compared to a Brownian motion with continuous paths. In empirical analysis of financial price data, one may detect big jumps, however, the small jumps are very hard to separate from the observations of a Brownian motion. Thus, it is not a simple task to decide whether a Lévy process with jumps or a Brownian motion is governing the small variations in a stock price, say.

In this paper we focus on the stability of Merton's problem with respect to model choice. In particular, we analyse what happens when the small jumps of the Lévy process driving the asset price dynamics is approximated by a Brownian motion. This would mimic a situation where we have two investors, one believing in a pure-jump Lévy process, and another which thinks the small variations in prices come from a Brownian motion. Asmussen and Rosinski [5] show that in fact the small jumps of a Lévy process has a central limit type behaviour towards a Brownian motion, which tells that one may empirically not be able to distinguish between two such models. The question is then to what extent this transfers over to the optimal portfolio selection problem. We pose the problem as an approximation of asset price models, where we either ignore or substitute jumps in the Lévy process smaller than a threshold ϵ. To substitute, we use a Brownian motion. Indeed, our analysis shows that the optimal investment in the risky asset is stable with respect to the different approximations. We are able to classify the convergence rate as being proportional to the variance of the small jump part of the Lévy process.

A general approach to stability of stochastic control problems are provided by Larsen and Žitković [46]. They investigate the influence of estimation errors in the parameters of the underlying financial assets. Jakobsen, Karlsen and La Chioma [43] are deriving stability results for the Hamilton-Jacobi-Bellman equation for stochastic control problems, and derive error estimates for approximative viscosity solutions. In a paper by Benth, Di Nunno and Khedher [9], stability for option pricing and hedging have been considered based on similar Lévy approximations as in the present paper. Here, the authors prove that prices and hedges converge at a rate given by the variance of the small jumps of the Lévy process, similar to our findings.

Our results are presented as follows. In Section 2 we state the control problems and recall some results on these. Afterwards we discuss in Section 3 how the approximation of the Lévy process influences the integral equation which gives implicitly the solution of the control problem. In Section 4 we study the convergence of the controls and derive convergence rates, which is illustrated by some numerical examples. The convergence of the value functions is treated in Section 5, and in Section 6 we analyse the wealth processes.

2.2 A review of Merton's portfolio optimization problem

We recall the Merton's portfolio optimization problem in the Lévy case with and without consumption, and review some relevant results from Benth et al. [13, 14].

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $\{\mathcal{F}_t\}_{t\geq 0}$ a given filtration satisfying the usual conditions. We consider a financial market consisting of a stock and a bond. Let the bond dynamics be given by

$$dB(t) = rB(t)dt\,,$$

where $r > 0$ is the constant interest rate. The value of the stock follows a process given by

$$S(t) = S(0)e^{\xi t + L(t)}$$

where ξ is a constant and L a pure jump Lévy process with Lévy-Khintchine decomposition

$$L(t) = \int_0^t \int_{|z|<1} z\tilde{N}(ds, dz) + \int_0^t \int_{|z|\geq 1} zN(dt, dz).$$

Here, $N(dt, dz)$ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R} \setminus \{0\}$ with intensity measure $dt \times \nu(dz)$, and $\nu(dz)$ being the Lévy measure, that is, a σ-finite Borel measure on $\mathbb{R} \setminus \{0\}$ with

$$\int_{\mathbb{R}\setminus\{0\}} \min(1, z^2)\, \nu(dz) < \infty\,.$$

We denote by $\tilde{N}(ds, dz) = N(dt, dz) - dt\nu(dz)$ the compensated Poisson random measure of N. In the sequel, we suppose that the Lévy process is exponentially integrable, that is, its Lévy measure satisfies the condition

$$\int_1^\infty e^{2z}\, \nu(dz) < \infty\,. \tag{2.2.1}$$

This will ensure that the stock price dynamics has finite expectation and variance, but also it will be necessary for the analysis to come in order to derive convergence rates for approximative portfolio strategies.

Consider an investor who puts her money in the stock and bond to optimize her utility. Let $\pi(t)$ denote the fraction of her wealth invested in the stock and $c(t)$ her rate of consumption at time t. The dynamics of the wealth $X^{(\pi,c)}$ becomes (see Benth et al. [13])

$$\begin{aligned} dX^{(\pi,c)}(t) &= X^{(\pi,c)}(t)(r+(\hat{\mu}-r)\pi(t))dt - c(t)\,dt \\ &\quad + X^{(\pi,c)}(t-)\pi(t-)\int_{\mathbb{R}\setminus\{0\}}(e^z-1)\tilde{N}(dt,dz), \end{aligned} \tag{2.2.2}$$

where

$$\hat{\mu} = \xi + \int_{\mathbb{R}\setminus\{0\}}(e^z - 1 - z1_{|z|<1})\nu(dz)$$

is the drift of the stock price dynamics and $X(t-)$ denotes the left-limit of $X(t)$. We denote by $X^{(\pi,c)}(0) = x$ the initial wealth of the investor, and assume $r < \hat{\mu}$. The last condition ensures that the stock gives a higher average return than the bond.

We define the set of *admissible controls* $\mathcal{A}_x$ to consist of those investment-consumption plans (π, c) such that

1. π is progressively measurable with values in $[0,1]$,
2. c is a positive and adapted process such that $\int_0^t \mathbb{E}[c(s)]ds < \infty$ for all $t \geq 0$,
3. c is such that $X^{(\pi,c)}(t) \geq 0$ almost everywhere for all $t \geq 0$.

We will restrict our attention to admissible controls, $(\pi, c) \in \mathcal{A}_x$. Observe that we constrain the invested fraction of wealth in the stock to be between 0 and 1, meaning that we cannot short sell stocks or borrow money to invest more than our wealth in stocks. One may extend the theory in Benth et al. [13] to $\pi \in [\underline{\pi}, \overline{\pi}]$, for $\underline{\pi} < 0$ and $\overline{\pi} > 0$. In the considerations to come, we can also include such a case with some additional effort.

The utility derived by the investor comes from consumption, and we suppose that she has a power utility function of HARA type, that is, $U(x) = x^\gamma/\gamma$ for a risk aversion parameter $\gamma \in (0,1)$. Letting $\delta > 0$ be a constant discount rate, the value function is defined by

$$V(x) = \sup_{c,\pi\in\mathcal{A}_x} \mathbb{E}^x\Big[\int_0^\infty e^{-\delta t}\big[\frac{c_t^\gamma}{\gamma}\big]dt\Big]\,. \tag{2.2.3}$$

By dynamic programming, the Hamilton-Jacobi-Bellman (HJB) equation takes the form

$$\begin{aligned} &\max_{c\geq 0,\pi\in[0,1]}\Big[(r+(\hat{\mu}-r)\pi_t)xv'(x) - cv'(x) - \delta v(x) + \frac{c^\gamma}{\gamma} \\ &\qquad + \int_{\mathbb{R}\setminus\{0\}}\big(v(x+\pi x(e^z-1)) - v(x) - \pi x v'(x)(e^z-1)\big)\nu(dz)\Big] = 0\,. \end{aligned} \tag{2.2.4}$$

Benth et al. [13] show that V is a viscosity solution of the HJB-equation. Moreover, the optimal investment strategy turns out to be a constant π^* solving implicitly the integral equation

$$\int_{\mathbb{R}\setminus\{0\}} \left([1+\pi(e^z-1)]^{\gamma-1}(e^z-1)-(e^z-1)\right)\nu(dz) = r-\hat{\mu}\,. \tag{2.2.5}$$

The optimal consumption is given as a constant rate of the wealth,

$$c^*(t) = X^{(\pi,c)}(t)\frac{1-\gamma}{\delta-k(\gamma)} \tag{2.2.6}$$

where

$$k(\gamma) = \gamma(r+(\hat{\mu}-r)\pi^*) + \int_{\mathbb{R}\setminus\{0\}} \left([1+\pi^*(e^z-1)]^{\gamma}-1-\gamma\pi(e^z-1)\right)\nu(dz).$$

We remark in passing that one may consider the simplified problem of maximizing terminal wealth only, and not consume anything from the portfolio. The value function becomes in this case

$$V(x) = \sup_{\pi\in\mathcal{A}_x} \mathbb{E}^x\left[\frac{1}{\gamma}X^\pi(T)^\gamma\right], \tag{2.2.7}$$

where we use the obvious definition of the set of admissible controls and the wealth process X^π (the latter is given by $X^{(\pi,0)}$). As it turns out, the optimal investment strategy is still a constant fraction of wealth placed in the stock, solving the integral equation (2.2.5).

We end this section with a discussion on conditions ensuring the existence and uniqueness of an optimal portfolio investment strategy $\pi^* \in [0,1]$. For this purpose, define

$$F(\pi) = \int_{\mathbb{R}\setminus\{0\}} \left([1+\pi(e^z-1)]^{\gamma-1}(e^z-1)-(e^z-1)\right)\nu(dz) + (\hat{\mu}-r)\,, \tag{2.2.8}$$

which is a continuous function on $[0,1]$ under our exponential integrability hypothesis on $\nu(dz)$. It holds

$$F(0) = \hat{\mu}-r\,,$$

which is positive by assumption on $\hat{\mu}$ and r. We have:

Lemma 2.2.1. *Assume that the Lévy measure and γ satisfy*

$$\int_{\mathbb{R}\setminus\{0\}} (e^z-1)(1-e^{-(1-\gamma)z})\,\nu(dz) > \hat{\mu}-r\,. \tag{2.2.9}$$

Then there exists a unique $\pi^* \in (0,1)$ solving (2.2.5).

Proof. By commuting differentiation and integration (see Folland [38]), we find for F in (2.2.8) that

$$F'(\pi) = -(1-\gamma)\int_{\mathbb{R}\setminus\{0\}} (e^z-1)^2[1+\pi(e^z-1)]^{\gamma-2}\,\nu(dz)\,.$$

Since $(\exp(z)-1)^2$ and $[1+\pi(\exp(z)-1)]^{\gamma-2}$ are both positive as long as $\pi\in[0,1]$, we find that $F'(\pi)<0$. Hence, F is strictly decreasing on $[0,1]$. Therefore, we have a unique solution $\pi^*\in(0,1)$ of (2.2.5) as long as $F(1)<0$. But this is ensured by the condition in the Lemma. $\square$

Note that $(\exp(z)-1)(1-\exp(-(1-\gamma)z))$ is positive for all $z \in \mathbb{R}$. Hence, the left-hand side of the condition (2.2.9) is positive. Hence, the condition therefore gives a relation between the Lévy measure and the risk aversion on one hand, and the excess return $\hat{\mu}-r$ on the other. Given the optimal control π^*, we have the optimal consumption process $c^*(t)$ as well.

For the analysis to come, it is convenient to introduce a function $f(\pi, z)$ defined as

$$f(\pi, z) = [1+\pi(e^z-1)]^{\gamma-1}(e^z-1) - (e^z-1)\,. \tag{2.2.10}$$

Furthermore, let $g(\pi)$ be

$$g(\pi) = \int_{\mathbb{R}\setminus\{0\}} f(\pi, z)\nu(dz)\,. \tag{2.2.11}$$

Then, from the definition of $F(\pi)$ we see that the integral equation (2.2.5) may be formulated compactly as

$$g(\pi) = r - \hat{\mu}. \tag{2.2.12}$$

We shall make use of these two functions when we move on in the next Section to consider approximations of the control problem of Merton.

2.3 The control problem with approximated driving process

In this section we examine the convergence properties of Merton's portfolio problem when we approximate the Lévy process L in the stock price dynamics. In particular, we consider two approximations, one where the small jumps of L are neglected, and another where we substitute the small jumps by a scaled Brownian motion. These two approximations will lead to different HJB-equations, and thus to different controls and value functions. We analyse the convergence to the original portfolio problem, and establish rates.

2.3.1 Approximating L by neglecting the small jumps

By appealing to the Lévy-Kintchine representation of L, we can write for a given $0<\epsilon<1$,

$$\begin{aligned} L(t) &= \int_0^t \int_{|z|<1} z\tilde{N}(ds,dz) + \int_0^t \int_{|z|\geq 1} zN(ds,dz) \qquad (2.3.1)\\ &= \int_0^t \int_{|z|<\epsilon} z\tilde{N}(ds,dz) + \int_0^t \int_{\epsilon<|z|<1} z\tilde{N}(ds,dz) + \int_0^t \int_{|z|\geq 1} zN(ds,dz). \end{aligned}$$

Introduce an approximation of L which neglects jumps smaller than ϵ:

$$L_{N,\epsilon}(t) = \int_0^t \int_{\epsilon<|z|<1} z\tilde{N}(ds,dz) + \int_0^t \int_{|z|\geq 1} zN(ds,dz).$$

Then, the Levy measure $\nu_{N,\epsilon}$ of $L_{N,\epsilon}$ is

$$\nu_{N,\epsilon}(dz) := \begin{cases} \nu(dz)\,, & |z| > \epsilon \\ 0\,, & \text{otherwise.} \end{cases}$$

The neglection of the small jumps influences the Levy measure and so indirectly also all terms which include a Lévy integral, as the drift of the stock price dynamics, the optimal controls and eventually the value function.

With obvious definition, we denote by $X_{N,\epsilon}^{(\pi,c)}$ the wealth process for admissible controls $(\pi, c) \in \mathcal{A}_{x,N,\epsilon}$. Furthermore, $V_{N,\epsilon}(x)$ denotes the value function.

Tracing through the derivation of Benth et al. [13] using $L_{N,\epsilon}$ in the stock price dynamics, leads to the following integral equation for the optimal control $\pi^*_{N,\epsilon}$

$$\int_{\mathbb{R}\setminus\{0\}} f(\pi_{N,\epsilon}, z)\nu_{N,\epsilon}(dz) = r - \hat{\mu}_{N,\epsilon} \tag{2.3.2}$$

with

$$\begin{aligned} \hat{\mu}_{N,\epsilon} &= \xi + \int_{|z|>\epsilon} (e^z - 1 - z1_{|z|<1})\,\nu(dz) \\ &= \hat{\mu} - \int_{|z|<\epsilon} (e^z - 1 - z)\,\nu(dz)\,. \end{aligned} \tag{2.3.3}$$

Furthermore, the optimal consumption process, denoted $c^*_{N,\epsilon}$ will be the same as for the non-approximated case, except that we insert the optimal control $\pi^*_{N,\epsilon}$ in (2.2.6) and use $\nu_{N,\epsilon}$ as the Lévy measure in the definition of k.

Let us investigate conditions for the existence of a unique solution to (2.3.2). Introduce the function $F_{N,\epsilon}(\pi)$ as

$$F_{N,\epsilon}(\pi) = \int_{\mathbb{R}\setminus\{0\}} \left([1+\pi(e^z-1)]^{\gamma-1}(e^z-1) - (e^z-1)\right)\nu_{N,\epsilon}(dz) + (\hat{\mu}_{N,\epsilon} - r)\,. \tag{2.3.4}$$

The optimal investment strategy $\pi^*_{N,\epsilon}$ is given as a root of the function $F_{N,\epsilon}$. Observe that the derivative of $F_{N,\epsilon}$ is negative, similar as to the case of $\epsilon = 0$. Thus, $F_{N,\epsilon}$ is a continuous function which is strictly decreasing. It will have a root in the interval $[0, 1]$ if and only if $F_{N,\epsilon}(0) > 0$ and $F_{N,\epsilon}(1) < 0$. But this is equivalent to

$$\hat{\mu} - r > \int_{|z|<\epsilon} (e^z - 1 - z)\,\nu(dz)$$

and

$$\int_{\mathbb{R}\setminus\{0\}} (e^z-1)(1-e^{-(1-\gamma)z})\,\nu_{N,\epsilon}(dz) < \hat{\mu} - r - \int_{|z|<\epsilon} (e^z - 1 - z)\,\nu(dz)\,.$$

We note that by Taylor expansion, the integral

$$\int_{|z|<\epsilon} (e^z - 1 - z)\,\nu(dz)$$

will be approximately equal to $\sigma^2(\epsilon)$ which tends to zero as $\epsilon \to 0$. Recalling that $\hat{\mu} > r$, we are ensured the existence and uniqueness of a solution $\pi^*_{N,\epsilon}$ by choosing ϵ sufficiently small if the condition (2.2.9) in Lemma 2.2.1 holds (that is, the condition for existence and uniqueness of $\pi^* \in [0,1]$).

To emphasize the difference of (2.3.2) from the original equation (2.2.5), reorganize to show that (2.3.2) is equivalent to

$$\int_{\mathbb{R}\setminus\{0\}} f(\pi_{N,\epsilon}, z)\nu(dz) - \int_{|z|<\epsilon} f(\pi_{N,\epsilon}, z)\,\nu(dz) = r - \left(\hat{\mu} - \int_{|z|<\epsilon} (e^z - 1 - z)\,\nu(dz)\right).$$

Or, using the function g in (2.2.11), we have

$$g(\pi_{N,\epsilon}) = r - \hat{\mu} + \int_{|z|<\epsilon} \left([1 + \pi_{N,\epsilon}(e^z - 1)]^{\gamma-1}(e^z - 1) - z\right)\nu(dz).$$

Introduce the function $h(\pi, z)$ for $|z| < 1$ by

$$h(\pi, z) := [1 + \pi(e^z - 1)]^{\gamma-1}(e^z - 1) - z\,. \tag{2.3.5}$$

Then, it finally follows that $\pi^*_{N,\epsilon}$ is the solution of the integral equation

$$g(\pi_{N,\epsilon}) = r - \hat{\mu} + \int_{|z|<\epsilon} h(\pi_{N,\epsilon}, z)\,\nu(dz). \tag{2.3.6}$$

In the analysis of convergence of $\pi^*_{N,\epsilon}$ to π^* as $\epsilon \to 0$, this representation of the optimal control is attractive.

2.3.2 Approximating L by substituting small jumps by Brownian motion

An alternative to truncating off the small jumps, is to approximate them by an appropriately scaled Brownian motion as discussed in Asmussen and Rosinski [5]. More precisely, we introduce the process

$$L_{W,\epsilon}(t) = \sigma(\epsilon)W(t) + L_{N,\epsilon}(t)\,, \tag{2.3.7}$$

where W is a Brownian motion (independent of L) and

$$\sigma^2(\epsilon) := \int_{|z|<\epsilon} z^2\nu(dz)\,, \tag{2.3.8}$$

is the variance of the small jumps (at least for symmetric Lévy processes). Note that $\sigma^2(\epsilon)$ is finite since $\nu(dz)$ integrates z^2 around the origin by definition. Moreover, by monotone convergence, it holds that

$$\lim_{\epsilon\to 0} \sigma^2(\epsilon) = 0\,.$$

It will be clear later that $\sigma^2(\epsilon)$ gives the rate of convergence in the approximations of the original portfolio optimization problem.

As in Benth et al. [14], an additional Brownian component does not change the general form of the solution of the control problem. We denote the wealth process by $X_{W,\epsilon}^{(\pi,c)}$ for admissible controls $(\pi, c) \in \mathcal{A}_{x,W,\epsilon}$, with an obvious definition of these. The value function in this case is denoted $V_{W,\epsilon}(x)$.

We can derive an integral equation for the optimal investment strategy, still being a constant $\pi_{W,\epsilon}^*$, but now solving the integral equation

$$\int_{\mathbb{R}\setminus\{0\}} f(\pi_{W,\epsilon}, z)\nu_{N,\epsilon}(dz) = r - \hat{\mu}_{W,\epsilon} + (1-\gamma)\sigma^2(\epsilon)\pi_{W,\epsilon}\,,$$

with

$$\begin{aligned}\hat{\mu}_{W,\epsilon} &= \xi + \int_{|z|>\epsilon} (e^z - 1 - z1_{|z|<1})\,\nu(dz) + \frac{1}{2}\sigma^2(\epsilon)\\ &= \hat{\mu} - \int_{|z|<\epsilon} (e^z - 1 - z)\,\nu(dz) + \frac{1}{2}\sigma^2(\epsilon).\end{aligned}$$

The optimal consumption process $c_{W,\epsilon}^*$ is given by

$$c_{W,\epsilon}^*(t) = X^{(\pi_{W,\epsilon}^*,c)}(t)\frac{1-\gamma}{\delta - k_{W,\epsilon}(\gamma)} \tag{2.3.9}$$

where

$$\begin{aligned}k_{W,\epsilon}(\gamma) = \gamma(r + (\hat{\mu}_{W,\epsilon} - r)\pi_{W,\epsilon}^*) - \frac{1}{2}\sigma^2(\epsilon)(\pi_{W,\epsilon}^*)^2\gamma(\gamma-1)\\ + \int_{\mathbb{R}\setminus\{0\}} \big([1+\pi_{W,\epsilon}^*(e^z-1)]^\gamma - 1 - \gamma\pi_{W,\epsilon}^*(e^z-1)\big)\nu_{N,\epsilon}(dz)\,.\end{aligned}$$

Again, we reformulate the equation for the optimal investment strategy in terms of g, in order to find

$$g(\pi_{W,\epsilon}) = r - \hat{\mu} + \sigma^2(\epsilon)\Big((1-\gamma)\pi_{W,\epsilon} - \frac{1}{2}\Big) + \int_{|z|<\epsilon} h(\pi_{W,\epsilon}, z)\,\nu(dz)\,. \tag{2.3.10}$$

Additionally to $\int_{|z|<\epsilon} h(\pi_{W,\epsilon}, z)\,\nu(dz)$, that also appeared in (2.3.6), we have a term with $\sigma^2(\epsilon)$ on the right hand side of (2.3.10).

We state conditions for the existence and uniqueness of an optimal strategy $\pi_{W,\epsilon}^* \in [0,1]$. Define the function $F_{W,\epsilon}(\pi)$ as

$$F_{W,\epsilon}(\pi) = \int_{\mathbb{R}\setminus\{0\}} \big([1+\pi(e^z-1)]^{\gamma-1}(e^z-1) - (e^z-1)\big)\,\nu_{N,\epsilon}(dz) + (\hat{\mu}_{W,\epsilon} - r) - (1-\gamma)\sigma^2(\epsilon)\pi\,.$$

Similar to the case of neglecting the small jumps, $F_{W,\epsilon}(\pi)$ is a strictly decreasing continuous function, which has a root in the interval $[0,1]$ if and only if $F_{W,\epsilon}(0) > 0$ and $F_{W,\epsilon}(1) < 0$. This is equivalent to

$$\hat{\mu} - r > \int_{|z|<\epsilon} (e^z - 1 - z)\,\nu(dz) - \frac{1}{2}\sigma^2(\epsilon)$$

and

$$\int_{\mathbb{R}\setminus\{0\}} (e^z - 1)(1 - e^{-(1-\gamma)z})\,\nu_{N,\epsilon}(dz) < \hat{\mu} - r - \int_{|z|<\epsilon} (e^z - 1 - z)\,\nu(dz) - (\frac{1}{2} - \gamma)\sigma^2(\epsilon)\,.$$

For the same reasons as before, we are ensured the existence and uniqueness of a solution $\pi^*_{W,\epsilon}$ by choosing ϵ sufficiently small if the condition (2.2.9) in Lemma 2.2.1 holds.

Let us discuss an example. In finance, the normal inverse Gaussian (NIG) distribution turns out to model the (log-)returns of financial asset prices very well. There exist several empirical studies where this distribution was applied, but we refer to Bølviken and Benth [26] which studied Norwegian stock prices. The NIG Lévy process is a pure-jump process, where the Lévy measure is explicitly known as

$$\nu(dz) = \frac{\alpha\delta}{\pi|z|} K_1(\alpha|z|)e^{\beta z}\,dz$$

for a NIG($\mu, \beta, \alpha, \delta$) process. Here, K_1 is the modified Bessel function of the third kind of index 1. In Rydberg [51] it was suggested to approximate the small jumps of the NIG process by a Brownian motion scaled by $\sigma(\epsilon)$, as discussed above. In Asmussen and Rosinski [5] this example was further elaborated, and they show that

$$\sigma^2(\epsilon) \sim \frac{2\delta}{\pi} \times \epsilon\,.$$

We remark that the NIG Lévy process was also used in Benth et al. [13] as a motivation for their studies of the Merton portfolio optimization problem for pure-jump Lévy processes.

2.4 Convergence rates for the optimal investment strategy

In this Section we prove that the approximative investment strategies $\pi^*_{N,\epsilon}$ and $\pi^*_{W,\epsilon}$ both converge to π^* as $\epsilon \to 0$. Moreover, we derive rates of convergence for both approximations in terms of the variance of the small jumps $\sigma^2(\epsilon)$.

2.4.1 Approximation of L by neglecting small jumps

Consider the case where we derive the optimal portfolio strategy $\pi^*_{N,\epsilon}$ based on an approximation where the small jumps are simply neglected. We have the following result.

Proposition 2.4.1. *The control $\pi^*_{N,\epsilon}$ solving* (2.3.2) *converges to the control π^* derived from* (2.2.5) *when $\epsilon \to 0$. In particular, it holds*

$$|\pi^*_{N,\epsilon} - \pi^*| \le C_N\sigma^2(\epsilon)\,,$$

for a constant $C_N > 0$ independent of ϵ.

Proof. Recall the definition of the function f in (2.2.10) to see that

$$\frac{\partial}{\partial\pi} f(\pi, z) := f'(\pi, z) = (\gamma - 1)(e^z - 1)^2[1 + \pi(e^z - 1)]^{\gamma-2}$$

is negative for all $z \in \mathbb{R}$. The risk aversion parameter γ is assumed to be between $(0,1)$. Hence, f is a continuous and strictly decreasing function of $\pi \in [0,1]$. Moreover, it is negative for all $\pi \in (0,1),\ z \in \mathbb{R}$ since $f(0) = 0$. Then, it follows that $g(\pi) = \int_{\mathbb{R}\setminus\{0\}} f(\pi,z)\nu(dz)$ in (2.2.11) is also strictly decreasing and negative. As

$$|f(\pi,z)| \le |f(1,z)| \text{ and } \int_{\mathbb{R}\setminus\{0\}} |f(1,z)|\nu(dz) < \infty\,,$$

the parameter-dependent integral defining $g(\pi)$ is continuous by Theorem 11.4 in Schilling [54]. Therefore, the inverse g^{-1} exists and is continuous on the image $g([0,1])$, and we can write the optimal control as

$$\pi^* = g^{-1}(r - \hat{\mu}).$$

Moreover, from elementary calculus the derivative of the inverse of a function $g(\pi) = y$ can be written as

$$\begin{aligned}(g^{-1})'(y) &= \frac{1}{g'(g^{-1}(y))} = \frac{1}{g'(\pi)}\\ &= \frac{1}{\frac{\partial}{\partial\pi}\int f(\pi,z)\,\nu(dz)} = \frac{1}{\int \frac{\partial}{\partial\pi} f(\pi,z)\,\nu(dz)}\end{aligned}$$

where we are allowed to commute integration and differentiation using Theorem 11.5 in Schilling [54] as long as $|\frac{\partial}{\partial\pi} f(\pi,z)| \le w(z)$ for $w(z)$ being an integrable function. But for $z > 0$ we find that

$$\left|\frac{\partial f}{\partial \pi}\right| \le (1-\gamma)(e^z - 1)^2$$

whereas for $z < 0$ we find

$$\left|\frac{\partial f}{\partial \pi}\right| \le (1-\gamma)(e^z - 1)^2 e^{-(2-\gamma)z}\,.$$

By the exponential integrability hypothesis on $\nu(dz)$, this defines an integrable function $w(z)$ verifying the commuting of integration of differentiation.

We continue with the proof of convergence. For $z > 0$, $\partial f/\partial\pi$ is monotonely increasing in $\pi \in [0,1]$ and for $z < 0$ it is monotonely decreasing. Additionally, $\partial f/\partial\pi$ is negative for all $z \in \mathbb{R}$. So,

$$\begin{aligned}\frac{\partial}{\partial\pi} f(\pi,z) &\le \frac{\partial}{\partial\pi} f(0,z),\ z < 0\\ \frac{\partial}{\partial\pi} f(\pi,z) &\le \frac{\partial}{\partial\pi} f(1,z),\ z > 0.\end{aligned}$$

Next we apply this to find an approximation for $(g^{-1})'(y)$:

$$\int_{\mathbb{R}\setminus\{0\}} \frac{\partial f}{\partial\pi}(\pi,z)\nu(dz) = \int_{\mathbb{R}^+} \frac{\partial f}{\partial\pi}(\pi,z)\nu(dz) + \int_{\mathbb{R}^-} \frac{\partial f}{\partial\pi}(\pi,z)\nu(dz)$$

$$\leq \int_{\mathbb{R}^+} \frac{\partial f}{\partial \pi}(1, z)\nu(dz) + \int_{\mathbb{R}^-} \frac{\partial f}{\partial \pi}(0, z)\nu(dz)$$

where we have

$$\frac{\partial f}{\partial \pi}(0, z) = (\gamma - 1)(e^z - 1)^2$$
$$\frac{\partial f}{\partial \pi}(1, z) = (\gamma - 1)(e^z - 1)^2 e^{z(\gamma-2)} .$$

Hence,

$$\begin{aligned}\Big| \int_{\mathbb{R}\setminus\{0\}} f'(\pi, z)\nu(dz)\Big| &\geq \Big| \int_{\mathbb{R}^+} f'(1, z)\nu(dz) + \int_{\mathbb{R}^-} f'(0, z)\nu(dz)\Big| \\ &= (1-\gamma)\left\{\int_0^\infty ((e^z - 1)^2 e^{-(2-\gamma)z})\,\nu(dz) + \int_{-\infty}^0 (e^z - 1)^2\,\nu(dz)\right\} \\ &=: L^{-1} .\end{aligned}$$

This implies that

$$|(g^{-1})'(y)| = \frac{1}{\big| \int_{\mathbb{R}\setminus\{0\}} f'(\pi, z)\nu(dz)\big|} \leq L . \tag{2.4.1}$$

With this bound on the derivative of the inverse function, we move on to estimate the error:

By applying the Mean Value Theorem in calculus to the equations (2.2.12) and (2.3.6), we find

$$\begin{aligned}|\pi^*_{N,\epsilon} - \pi^*| &= \Big|g^{-1}\Big(r - \hat{\mu} + \int_{|z|<\epsilon} h(\pi^*_{N,\epsilon}, z)\nu(dz)\Big) - g^{-1}(r - \hat{\mu})\Big| \\ &= |(g^{-1})'(\theta)|\Big|r - \hat{\mu} + \int_{|z|<\epsilon} h(\pi^*_{N,\epsilon}, z)\nu(dz) - (r - \hat{\mu})\Big| \\ &\leq L\Big| \int_{|z|<\epsilon} h(\pi^*_{N,\epsilon}, z)\nu(dz)\Big|\end{aligned}$$

for some

$$\theta \in [\min\{r - \hat{\mu}, r - \hat{\mu} + \int_{|z|<\epsilon} h(\pi^*_{N,\epsilon}, z)\,\nu(dz)\}, \max\{r - \hat{\mu}, r - \hat{\mu} + \int_{|z|<\epsilon} h(\pi^*_{N,\epsilon}, z)\,\nu(dz)\}] .$$

Next, let us estimate the term involving h. Since $h(\pi, z)$ is decreasing in $\pi \in [0, 1]$, we have

$$h(0, z) \geq h(\pi, z) \geq h(1, z) .$$

Note that h can be both positive and negative, which makes it difficult for estimations of the absolute value of the integral of h. Divide the domain of definition of h in z into those parts where h is positive and negative for fixed π and γ:

$$A_\pi := \{z \in \mathbb{R} : h(\pi, z) \geq 0\} ,$$
$$B_\pi := \{z \in \mathbb{R} : h(\pi, z) < 0\} .$$

Then, for $z \in A_\pi$, we find from Taylor expansions

$$|h(\pi, z)| \leq |h(0, z)| = |e^z - 1 - z| \leq \sum_{n=2}^{\infty} \frac{|z|^n}{n!} \leq z^2 e^{|z|},$$

and for $z \in B_\pi$ we have

$$\begin{aligned}|h(\pi, z)| &\leq |h(1, z)| = |e^{z(\gamma-1)}(e^z - 1) - z| = |e^{\gamma z} - e^{z(\gamma-1)} - z| \\ &\leq \sum_{n=2}^{\infty} \frac{|\gamma^n - (\gamma-1)^n|}{n!}|z|^n \leq z^2 \sum_{n=0}^{\infty} \frac{|\gamma^{n+2} - (\gamma-1)^{n+2}|}{n!}|z|^n \leq z^2 e^{|z|},\end{aligned}$$

as $|\gamma^{n+2} - (\gamma-1)^{n+2}| \leq 1$. It follows

$$\begin{aligned}|\pi^*_{N,\epsilon} - \pi^*| &\leq L\Big|\int_{|z|<\epsilon} h(\pi^*_{N,\epsilon}, z)\,\nu(dz)\Big| \\ &\leq L\Big(\Big|\int_{\{|z|<\epsilon\}\cap A_{\pi^*_{N,\epsilon}}} h(\pi^*_{N,\epsilon}, z)\,\nu(dz)\Big| + \Big|\int_{\{|z|<\epsilon\}\cap B_{\pi^*_{N,\epsilon}}} h(\pi^*_{N,\epsilon}, z)\,\nu(dz)\Big|\Big) \\ &\leq L\Big(\int_{\{|z|<\epsilon\}\cap A_{\pi^*_{N,\epsilon}}} |h(\pi^*_{N,\epsilon}, z)|\,\nu(dz) + \int_{\{|z|<\epsilon\}\cap B_{\pi^*_{N,\epsilon}}} |h(\pi^*_{N,\epsilon}, z)|\,\nu(dz)\Big) \\ &\leq Le^\epsilon\Big(\int_{\{|z|<\epsilon\}\cap A_{\pi^*_{N,\epsilon}}} z^2\,\nu(dz) + \int_{\{|z|<\epsilon\}\cap B_{\pi^*_{N,\epsilon}}} z^2\nu(dz)\Big) \\ &\leq Le^\epsilon \sigma^2(\epsilon) \\ &\leq Le^1 \sigma^2(\epsilon)\end{aligned}$$

as $\epsilon \leq 1$. This completes the proof. □

Remark that since $h(1, z)$ is increasing in $z \in [-1, 1]$, we find that

$$h(1, z) \geq h(1, -1) = e^{-\gamma}(1 - e) + 1\,.$$

It follows that $h(\pi, z)$ is positive for all $z \in \mathbb{R}$ and $\pi \in [0, 1]$ if

$$\gamma > \ln(e - 1) \approx 0.541. \tag{2.4.2}$$

Thus, if $\gamma \geq \ln(e - 1)$, we have

$$|h(\pi, z)| \leq |h(0, z)| = e^z - 1 - z\,.$$

This implies that

$$\Big|\int_{\mathbb{R}_0} h(\pi, z)\,\nu(dz)\Big| \leq e^\epsilon \sigma^2(\epsilon)$$

which simplifies the proof above.

In practice one may also be interested in a lower bound for the error. In the case (2.4.2), that is $h(\pi, z)$ is positive, it is relatively simple to find a lower bound, which turns out to be again proportional to the variance of the small jumps.

Corollary 2.4.2. *For $\gamma > \ln(e-1)$ it holds that*

$$|\pi^*_{N,\epsilon} - \pi^*| \geq \widetilde{C}_N \sigma^2(\epsilon)\,,$$

for a constant $\widetilde{C}_N > 0$ independent of ϵ.

Proof. From the proof of Proposition 2.4.1 we have that

$$|\pi^*_{N,\epsilon} - \pi^*| = |g^{-1}(\theta)|\Big|\int_{|z|<\epsilon} h(\pi^*_{N,\epsilon}, z)\nu(dz)\Big|$$

for some

$$\theta \in \Big[\min\{r - \hat{\mu}, r - \hat{\mu} + \int_{|z|<\epsilon} h(\pi^*_{N,\epsilon}, z)\,\nu(dz)\},\\ \max\{r - \hat{\mu}, r - \hat{\mu} + \int_{|z|<\epsilon} h(\pi^*_{N,\epsilon}, z)\,\nu(dz)\}\Big]\,.$$

The argumentation leading to an upper bound for $|g^{-1}(\theta)|$ in (2.4.1) leads analogously to a lower bound

$$|g^{-1}(\theta)| \geq L_2$$

with

$$L_2^{-1} := \Big|\int_{\mathbb{R}^+} f'(0,z)\nu(dz) + \int_{\mathbb{R}^-} f'(1,z)\nu(dz)\Big|.$$

As $h(\pi, z)$ is positive for $\gamma > \ln(e-1)$, decreasing in $\pi \in [0,1]$ and increasing in $|z| \leq 1$, we find

$$h(\pi, z) \geq h(1,z) \geq h(1,-1) \geq z^2 h(1,-1)$$

and

$$\Big|\int_{|z|<\epsilon} h(\pi^*_{N,\epsilon}, z)\nu(dz)\Big| \geq h(1,-1)\int_{|z|<\epsilon} z^2\nu(dz).$$

Hence,

$$|\pi^*_{N,\epsilon} - \pi^*| \geq L_2 h(1,-1)\sigma^2(\epsilon)$$

and the result follows. □

The Corollary yields that the error is proportional from below and above by $\sigma^2(\epsilon)$. Going back to the case of an NIG Lévy process L, then by neglecting the small jumps and solving the portfolio optimization problem would yield an error which could be bounded above as

$$|\pi^*_{N,\epsilon} - \pi^*| \leq C \times \epsilon\,.$$

But, the Corollary above tells us that the lower bound for the error is also proportional in ϵ.

2.4.2 Approximation of L by substituting small jumps by Brownian motion

We move on showing that the approximation using a Brownian motion leads to a convergence of $\pi^*_{W,\epsilon}$ to π^* with the same rate as for the case where small jumps are neglected. We formulate the result as a proposition:

Proposition 2.4.3. *The control $\pi^*_{W,\epsilon}$ solving* (2.3.10) *converges to the control π^* derived from* (2.2.12) *when $\epsilon \to 0$. In particular, it holds*

$$|\pi^*_{W,\epsilon} - \pi^*| \leq C_W \sigma^2(\epsilon)\,,$$

for a constant $C_W > 0$ independent of ϵ.

Proof. The proof follows the same line of arguments as in the proof of Prop. 2.4.1.

$$\begin{aligned}
|\pi^*_{W,\epsilon} - \pi^*| &= \big|g^{-1}\big(r - \hat{\mu} + \sigma^2(\epsilon)\big((1-\gamma)\pi^*_{W,\epsilon} - \frac{1}{2}\big) + \int_{|z|<\epsilon} h(\pi^*_{W,\epsilon}, z)\,\nu(dz)\big) \\
&\qquad - g^{-1}(r - \hat{\mu})\big| \\
&\leq L\big|r - \hat{\mu} + \sigma^2(\epsilon)\big((1-\gamma)\pi^*_{W,\epsilon} - \frac{1}{2}\big) - (r - \hat{\mu})\big| + \big|\int_{|z|<\epsilon} h(\pi^*_{W,\epsilon}, z)\,\nu(dz)\big| \\
&\leq L\Big(\sigma^2(\epsilon)\big(\frac{3}{2} - \gamma\big) + \big|\int_{|z|<\epsilon} h(\pi^*_{W,\epsilon}, z)\nu(dz)\big|\Big)\,.
\end{aligned}$$

Invoking the estimations on h from the proof of Prop. 2.4.1 gives the result. □

Inspecting the proofs of Prop. 2.4.1 and 2.4.3 shows that the constant in the convergence rate when neglecting the small jumps is given by $C_N = Le$, whereas for the Brownian motion approximation it is $C_W = L(3/2 - \gamma) + C_N > C_N$ as $\gamma \in (0, 1)$. Thus, the error estimate is in fact slightly worse when we use an approximation which gives a Lévy process with approximately the same variance, compared to an approximation where some of the noise is removed.

2.4.3 Improvement of the convergence rate when the small jumps are approximated by a Brownian motion

In fact, the triangle inequality in the proof of Proposition 2.4.3 is applied too early. The additional term resulting from the Brownian approximation chancels with the second order terms of the Taylor expansion of the function h in the proof of Proposition 2.4.3. We can therefore improve the rate of convergence in the case that the small jumps are approximated by a Brownian motion as the following Proposition shows

Proposition 2.4.4. *The control $\pi^*_{W,\epsilon}$ solving* (2.3.10) *converges to the control π^* derived from* (2.2.12) *when $\epsilon \to 0$. In particular, it holds*

$$|\pi^*_{W,\epsilon} - \pi^*| \leq C_W \epsilon \sigma^2(\epsilon)\,,$$

for a constant $C_W > 0$ independent of ϵ.

Proof. Applying the mean value theorem as in the proof of Proposition 2.4.1 results in

$$|\pi^*_{W,\epsilon} - \pi^*| \leq L\Big|\Big(\sigma^2(\epsilon)\big((1-\gamma)\pi^*_{W,\epsilon} - \frac{1}{2}\big) + \int_{|z|<\epsilon} h(\pi^*_{W,\epsilon}, z)\nu(dz)\Big)\Big|,$$

where the function $h(\pi, z)$ for $|z| < 1$ is given by

$$h(\pi, z) := [1 + \pi(e^z - 1)]^{\gamma-1}(e^z - 1) - z\,.$$

Now we don't use the triangle inequality as in the proof of Proposition 2.4.3, but do a Taylor expansion around $z = 0$:

$$\begin{aligned} h(\pi, z) &= -\big((1-\gamma)\pi^*_{W,\epsilon} - \frac{1}{2}\big)z^2 + \sum_{k=3}^{\infty} \frac{h^{(k)}(0,\pi)}{k!} z^k \\ &= -\big((1-\gamma)\pi^*_{W,\epsilon} - \frac{1}{2}\big)z^2 + z^3 \sum_{k=0}^{\infty} \frac{h^{(k+3)}(0,\pi)}{(k+3)!} z^k\,, \end{aligned}$$

and

$$\sum_{k=0}^{\infty} \frac{h^{(k+3)}(0,\pi)}{(k+3)!} z^k = \frac{h(\pi, z) + \big((1-\gamma)\pi^*_{W,\epsilon} - \frac{1}{2}\big)z^2}{z^3}\,. \tag{2.4.3}$$

With L'Hospitals rule it follows

$$\lim_{z\to 0} \frac{h(\pi, z) + \big((1-\gamma)\pi^*_{W,\epsilon} - \frac{1}{2}\big)z^2}{z^3} = \frac{1}{2}(\gamma-1)(\gamma-2)\pi^2 + \frac{3}{2}(\gamma-1)\pi + \frac{1}{6}\,.$$

Then, the fraction in (2.4.3) is bounded (in absolute values) by a constant c for all $z \in [-1, 1]$ and $\pi \in [0, 1]$ as $h(\pi, z)$ is bounded. We have then

$$\begin{aligned} |\pi^*_{W,\epsilon} - \pi^*| &\leq Lc \int_{|z|<\epsilon} |z|^3 \nu(dz) \\ &\leq Lc\, \epsilon\, \sigma^2(\epsilon)\,. \end{aligned}$$

□

2.4.4 Examples

First, let us assume the driving process L is a Poisson process N, compensated by its jump intensity λ,

$$L(t) = N(t) - \lambda t\,.$$

This is admittedly not a process which has "small jumps", since all jumps are of constant size 1. However, we would like to look at an example where we perturb this process by adding a Brownian motion component, that is

$$L_\epsilon(t) = L(t) + \epsilon W(t)\,.$$

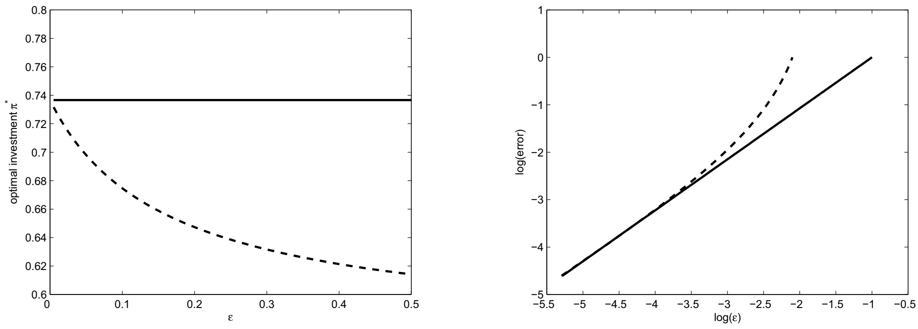

Figure 2.1: Error introduced by the Brownian motion term, when L follows a compensated Poisson process

This is not fitting into our analysis above, but the case here is to see the effect of perturbing the process L in a simple setting where much is known analytically. It serves as a "non-example" which is still relevant for our considerations.

We find the optimal control π^* by solving (2.2.12), which in this case becomes

$$\big([1+\pi^*(e-1)]^{\gamma-1}(e-1)-(e-1)\big)\lambda = r-((\gamma-\lambda)+(e-1)\lambda)\,.$$

This yields the solution

$$\pi^* = \left(\left(\frac{r-\gamma+\lambda}{(e-1)\lambda}\right)^{\frac{1}{\gamma-1}}-1\right)\frac{1}{e-1}.$$

The corresponding equation for solving π^*_ϵ, the optimal fraction to invest in the stock for the driving process L_ϵ, is

$$\begin{aligned}\big([1+\pi^*_\epsilon(e-1)]^{\gamma-1}(e-1)-(e-1)\big)\lambda = r &- ((\gamma-\lambda)+(e-1)\lambda) \\ &+((1-\gamma)\pi^*_\epsilon-\frac{1}{2})\epsilon\,,\end{aligned}$$

which we solve numerically.

To be concrete, we suppose an annual interest rate $r = 4.5\%$ and a jump intensity given by $\lambda = 0.5$, corresponding to an asset that jumps on average every second day. Furthermore, the risk aversion is set equal to $\gamma = 0.5$. This yields an optimal investment in the stock of $\pi^* = 0.7367$, that is, 73.67% of the wealth should go into the stock. In Figure 2.1 we plot the "exact solution" π^* against the approximative π^*_ϵ for ϵ ranging from 0.01 to 1. The approximative investment strategy is found by solving the equation using the built-in Matlab routine `fzero`. Since the approximative model has more noise/uncertainty, it naturally leads to an investment strategy less than π^*. Noteworthy is that the difference is rather big even for small ϵ's. For example, if $\epsilon = 0.1$, we have a relative error of approximately 5.3%, whereas for $\epsilon = 0.01$ it is 0.7%.

We see from the figure where we plot the error on log-scale against a log-epsilon that it corresponds very well to the line $-1 + 0.93 \times \log(\epsilon)$, which means that the error goes approximately as $C \times \epsilon^{0.93}$, slightly worse than a linear convergence in ϵ.

We now move on to consider the more interesting case of an NIG Lévy process and the approximation of such. We restrict our attention to the situation where we neglect the small jumps, and investigate the deviation between the "correct" portfolio strategy π^* and the approximative $\pi^*_{N,\epsilon}$.

Let the parameters be $\mu = 0$, $\beta = 0$, $\alpha = 50$ and $\delta = 0.03$ on a daily scale. The choices of α and δ seem to be natural estimates of stocks, see Bølviken and Benth [26]. A β with value zero corresponds to an assumption of symmetric logreturns, which is close to empirical findings as well (see again Bølviken and Benth [26] for examples). Furthermore, we suppose $r = 0.04/250$ and $\xi = 0.02/250$. The interest rate is therefore 4% while logreturns have a mean of 2%, measured annually, when we assume there are 250 trading days in a year. The compound interest rate $\hat{\mu}$ on the average stock price becomes 0.095, or 9.5%, annually. This is clearly above r, implying that the condition $\hat{\mu} - r > 0$ is satisfied. The risk aversion coefficient is supposed to be $\gamma = 0.5$.

We computed the optimal π^* using $\epsilon = 10^{-10}$ to avoid the singularity at the center of the Lévy measure of the NIG. The resulting optimal investment strategy became $\pi^* = 74\%$. Since the Lévy measure $\nu(dz)$ of an NIG Lévy process is supported on the whole real line (except zero, of course), we must truncate the integral in the expressions of $F_{N,\epsilon}(\pi)$ and $\hat{\mu}_{N,\epsilon}$, found in (2.3.4) and (2.3.3), resp. Integration over the domain $|z| > 1$ in these two expressions gave a contribution of the magnitude less than 10^{-20}, whereas the integral around the center in 0 has values on the level 10^{-2}. Thus, we truncate the integrals at ± 1, and consider only $|z| \leq 1$. Integration was performed in Matlab using the routine `quad`, which is an adaptive Simpson quadrature method.

We investigated the investment strategy $\pi^*_{N,\epsilon}$ when letting the ϵ range from 0.0005 to 0.007. The resulting strategies are plotted in Fig. 2.2 as asterisks, with the straight line being the optimal limiting case $\pi^* = 0.74$. We see that the optimal strategies when ignoring the small jumps are already approximately 0.07 lower in value than π^* for $\epsilon = 0.007$, meaning around 9.5% lower fraction invested in the risky asset. The truncated strategies are lower than 0.74, which may seem surprising at first sight since we have less risk as a result of ignoring small jumps. However, as it turns out, the small jumps will contribute significantly to the expected return of the asset. In fact, the expected return from $\epsilon = 0.007$ will be only 7.87%, compared with 9.5% for the case of no truncation. Hence, less noise is outweighed by reduced return in the stock, and the investor goes for a reduced position in the risky asset. Furthermore, from Fig. 2.2 we see that the error between the approximative optimal investment strategy and $\pi^* = 0.74$ goes close to linearity in ϵ. Recall that for the NIG, $\sigma^2(\epsilon)$ is proportional to ϵ.

2.5 Convergence of the value functions

We have seen that the controls $\pi^*_{N,\epsilon}$ and $\pi^*_{W,\epsilon}$ converge to the control π^* as $\epsilon \to 0$. We next investigate the convergence of the corresponding value functions. We first discuss maximization of terminal wealth, and then move on to analyse the convergence of the optimal consumption and the value functions.

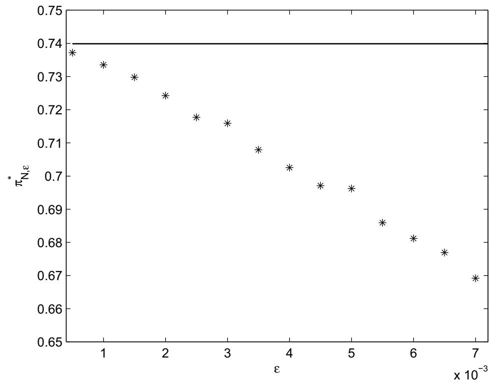

Figure 2.2: Plot of $\pi^*_{N,\epsilon}$ (asterisks) vs. the optimal π^* (complete line) for an NIG Lévy process L.

2.5.1 Maximising expected utility of terminal wealth

We consider the problem of optimizing terminal wealth, which means there is no consumption involved in the control problem. We find the following for the case of truncation of the small jumps:

Proposition 2.5.1. *It holds for every $x \in \mathbb{R}_+$,*

$$\lim_{\epsilon \to 0} V_{N,\epsilon}(x) = V(x).$$

Proof. Recalling the wealth process X^π with no consumption, we find that it is a geometric jump diffusion process with constant coefficients and with solution

$$\begin{aligned} X^{\pi^*}(t) = x \exp \big\{ & t(r + (\hat{\mu} - r)\pi) \\ & + t \int_{\mathbb{R}\setminus\{0\}} \big(\ln(1 + \pi^*(e^z - 1)) - \pi^*(e^z - 1) 1_{|z|<1} \big) \, \nu(dz) \\ & + \int_0^t \int_{\mathbb{R}\setminus\{0\}} \ln(1 + \pi^*(e^z - 1)) \tilde{N}(ds, dz) \big\}. \end{aligned}$$

Applying the formula in Ex. 1.6 in Øksendal and Sulem [58] , we find

$$\begin{aligned} V(x) &= \mathbb{E}[U(X^{\pi^*}(T))] = \mathbb{E}\left[\frac{1}{\gamma} (X^{\pi^*}(T))^\gamma \right] \\ &= \frac{1}{\gamma} x^\gamma \exp \big\{ \gamma T (r + (\hat{\mu} - r)\pi^*) \\ &\qquad + T \int_{\mathbb{R}\setminus\{0\}} \big([1 + \pi^*(e^z - 1)]^\gamma - 1 - \gamma \pi^*(e^z - 1) 1_{|z|<1} \big) \, \nu(dz) \big\} \end{aligned}$$

If small jumps are neglected, we find similarly

$$V_{N,\epsilon}(x) = \mathbb{E}[U(X^*_{N,\epsilon}(T))]$$

$$
\begin{aligned}
&= \frac{1}{\gamma} x^\gamma \exp \big\{ \gamma T (r + (\hat{\mu}_{N,\epsilon} - r)\pi^*_{N,\epsilon}) \\
&\qquad + T \int_{\mathbb{R}\setminus\{0\}} \big([1 + \pi^*_{N,\epsilon}(e^z - 1)]^\gamma - 1 - \gamma \pi^*_{N,\epsilon}(e^z - 1) 1_{|z|<1} \big) \, \nu_{N,\epsilon}(dz) \big\}.
\end{aligned} \tag{2.5.1}
$$

We have for $|z| < 1$

$$
|[1 + \pi^*_{N,\epsilon}(e^z - 1)]^\gamma - 1 - \gamma \pi^*_{N,\epsilon}(e^z - 1)| \le |e^{\gamma z} - 1 - \gamma(e^z - 1)| \,, \tag{2.5.2}
$$

and for $|z| > 1$,

$$
|[1 + \pi^*_{N,\epsilon}(e^z - 1)]^\gamma - 1| \le |e^{\gamma z} - 1| \,.
$$

Both estimates are integrable on their definition area. As $\pi^*_{N,\epsilon}$ converges to π^* and $\hat{\mu}_{N,\epsilon}$ to $\hat{\mu}$, the proposition follows with Lebesgue's convergence theorem. □

For the case of a Brownian approximation of the truncated small jumps we have:

Proposition 2.5.2. *It holds for every $x \in \mathbb{R}_+$,*

$$
\lim_{\epsilon \to 0} V_{W,\epsilon}(x) = V(x).
$$

Proof. A Brownian motion approximation of the small jumps as in Section 2.3.2 results in an additional Brownian component and $\sigma(\epsilon)$-terms in the wealth process:

$$
\begin{aligned}
X^*_{W,\epsilon}(t) = x \exp \big\{ & t \big[r + (\hat{\mu}_{W,\epsilon} - r)\pi_{W,\epsilon} - \frac{1}{2}\sigma^2(\epsilon)\pi^2_{W,\epsilon} \big] + \frac{1}{2}\sigma(\epsilon)\pi_{W,\epsilon} B_t \\
&+ t \int_{\mathbb{R}\setminus\{0\}} \big(\ln(1 + \pi^*_{W,\epsilon}(e^z - 1)) - \pi^*_{W,\epsilon}(e^z - 1) 1_{|z|<1} \big) \, \nu_{N,\epsilon}(dz) \\
&+ \int_0^t \int_{\mathbb{R}\setminus\{0\}} \ln(1 + \pi^*_{W,\epsilon}(e^z - 1)) \tilde{N}_\epsilon(ds, dz) \big\}.
\end{aligned}
$$

The value function becomes then

$$
\begin{aligned}
V_{W,\epsilon}(x) &= \mathbb{E}[U(X^*_{W,\epsilon}(T))] \\
&= \frac{1}{\gamma} x^\gamma \exp \big\{ \gamma T (r + (\hat{\mu}_{W,\epsilon} - r)\pi^*_{W,\epsilon} - \frac{1}{2}\sigma^2(\epsilon)\pi^2_{W,\epsilon}) \\
&\qquad + T \int_{\mathbb{R}\setminus\{0\}} \big([1 + \pi^*_{W,\epsilon}(e^z - 1)]^\gamma - 1 - \gamma \pi^*_{W,\epsilon}(e^z - 1) 1_{|z|<1} \big) \, \nu_{N,\epsilon}(dz) \big\} \,.
\end{aligned}
$$

Hence, as for the case of $V_{N,\epsilon}$ above, $V_{W,\epsilon}(x)$ converges to $V(x)$. □

2.5.2 Maximising expected utility of consumption

As we noted in Section 2, maximising wealth over optimal investment and consumption pairs (π, c) results in the same optimal strategy π^* as when maximising expected utility over terminal wealth. As our results show, approximations of these control problems leads to convergence of the optimal investment strategies, as well as the value functions for maximization of the utility of terminal wealth. We show next that including consumption does not alter these conclusions.

Proposition 2.5.3. *Let the value function be of the form (2.2.3) and suppose that the discount factor δ satisfies*

$$\delta > \gamma\hat{\mu} + \int_1^\infty (e^{\gamma z} - 1)\, \nu(dz)\,.$$

Then it holds for every $x \in \mathbb{R}_+$,

$$\lim_{\epsilon\to 0} V_{N,\epsilon}(x) = V(x)\,.$$

Proof. As the optimal consumption is a constant fraction of wealth, the wealth process is again a geometric jump diffusion process with constant coefficients and with solution

$$\begin{aligned} X^{\pi^*}(t) = x \exp\Big\{ & t(r + (\hat{\mu} - r)\pi^* - c^*) \\ & + t\int_{\mathbb{R}\setminus\{0\}} \big(\ln(1 + \pi^*(e^z - 1)) - \pi^*(e^z - 1)1_{|z|<1}\big)\, \nu(dz) \\ & + \int_0^t \int_{\mathbb{R}\setminus\{0\}} \ln(1 + \pi^*(e^z - 1))\tilde{N}(ds, dz)\Big\} \end{aligned}$$

where

$$c^* = \frac{1-\gamma}{\delta - k(\gamma)}.$$

Then the value function takes the form

$$\begin{aligned} V(x) &= \mathbb{E}\Big[\int_0^\infty e^{-\delta t}\frac{(c^*(t))^\gamma}{\gamma} dt\Big] \\ &= \frac{1}{\gamma}(c^*)^\gamma x^\gamma \int_0^\infty \exp\Big\{t[-\delta + \gamma(r + (\hat{\mu} - r)\pi^* - c^*) \\ &\quad + \int_{\mathbb{R}\setminus\{0\}} \big([1 + \pi^*(e^z - 1)]^\gamma - 1 - \gamma\pi^*(e^z - 1)1_{|z|<1}\big)\nu(dz)]\Big\}dt \end{aligned}$$

where we have exchanged integration and expectation by appealing to Fubini's theorem and formula in Ex. 1.6 in Øksendal and Sulem [58]. For the value function to be finite, we must require that,

$$\begin{aligned} & -\delta + \gamma(r + (\hat{\mu} - r)\pi^* - c^*) \\ & \quad + \int_{\mathbb{R}\setminus\{0\}} \big([1 + \pi^*(e^z - 1)]^\gamma - 1 - \gamma\pi^*(e^z - 1)1_{|z|<1}\big)\, \nu(dz) < 0\,. \end{aligned} \tag{2.5.3}$$

Condition (2.5.3) depends on π^* and c^*. With π^* bounded in $[0,1]$ and c^* positive we find

$$\gamma(r + (\hat{\mu} - r)\pi^* - c^*) \leq \gamma\hat{\mu}.$$

Furthermore, the integrand in (2.5.3) is positive and increasing in π for $z > 1$ and negative and decreasing in π for $z < 1$. Then

$$\int_{\mathbb{R}\setminus\{0\}} \left([1 + \pi^*(e^z - 1)]^\gamma - 1 - \gamma\pi^*(e^z - 1)1_{|z|<1}\right) \nu(dz) < \int_1^\infty (e^{\gamma z} - 1)\nu(dz)$$

and condition (2.5.3) is fulfilled by the assumption on δ. Thus,

$$\begin{aligned} V(x) = &-\frac{1}{\gamma}(c^*)^\gamma x^\gamma[-\delta + \gamma(r + (\hat{\mu} - r)\pi^* - c^*) \\ &+ \int_{\mathbb{R}\setminus\{0\}} \left([1 + \pi^*(e^z - 1)]^\gamma - 1 - \gamma\pi^*(e^z - 1)1_{|z|<1}\right) \nu(dz)]^{-1}. \end{aligned}$$

Neglecting small jumps, we find similarly

$$\begin{aligned} V_{N,\epsilon}(x) &= \frac{1}{\gamma}(c^*_{N,\epsilon})^\gamma \int_0^\infty x^\gamma \exp\big\{t[-\delta + \gamma(r + (\hat{\mu}_{N,\epsilon} - r)\pi^*_{N,\epsilon} - c^*_{N,\epsilon}) \\ &\quad + \int_{\mathbb{R}\setminus\{0\}} \left([1 + \pi^*_{N,\epsilon}(e^z - 1)]^\gamma - 1 - \gamma\pi^*_{N,\epsilon}(e^z - 1)1_{|z|<1}\right)\nu_{N,\epsilon}(dz)]\big\}dt \\ &= -\frac{1}{\gamma}(c^*_{N,\epsilon})^\gamma x^\gamma[-\delta + \gamma(r + (\hat{\mu}_{N,\epsilon} - r)\pi^*_{N,\epsilon} - c^*_{N,\epsilon}) \\ &\quad + \int_{\mathbb{R}\setminus\{0\}} \left([1 + \pi^*_{N,\epsilon}(e^z - 1)]^\gamma - 1 - \gamma\pi^*_{N,\epsilon}(e^z - 1)1_{|z|<1}\right)\nu_{N,\epsilon}(dz)]^{-1} \end{aligned}$$

with

$$c^*_{N,\epsilon} = \frac{1 - \gamma}{\delta - k_{N,\epsilon}(\gamma)}$$

and

$$\begin{aligned} k_{N,\epsilon}(\gamma) = &\gamma(r + (\hat{\mu}_{N,\epsilon} - r)\pi^*_{N,\epsilon}) \\ &+ \int_{\mathbb{R}\setminus\{0\}} \left([1 + \pi^*_{N,\epsilon}(e^z - 1)]^\gamma - 1 - \gamma\pi^*_{N,\epsilon}(e^z - 1)\right)\nu_{N,\epsilon}(dz). \end{aligned}$$

$k_{N,\epsilon}(\gamma)$ converges to $k(\gamma)$ as $\hat{\mu}_{N,\epsilon}$ and $\pi^*_{N,\epsilon}$ converges and by appealing to Lebesgue's convergence theorem using estimate (2.5.2). The integral

$$\int_{\mathbb{R}\setminus\{0\}} \left([1 + \pi^*_{N,\epsilon}(e^z - 1)]^\gamma - 1 - \gamma\pi^*_{N,\epsilon}(e^z - 1)1_{|z|<1}\right)\nu_{N,\epsilon}(dz)$$

did already appear in (2.5.1) in the value function connected to maximising terminal wealth, where its convergence was discussed. Then the proposition follows. □

Approximating the truncated jumps by a Brownian motion preserves convergence of the value function also in the consumption case, as the next result shows:

Proposition 2.5.4. *Let the value function be of the form (2.2.3). Then it holds for every* $x \in \mathbb{R}_+$,

$$\lim_{\epsilon \to 0} V_{W,\epsilon}(x) = V(x).$$

Proof. In this case the value function has the form

$$V_{W,\epsilon}(x) = \frac{1}{\gamma}(c^*_{W,\epsilon})^\gamma \int_0^\infty x^\gamma \exp\Big\{t[-\delta + \gamma(r + (\hat{\mu}_{W,\epsilon} - r)\pi^*_{W,\epsilon} - \frac{1}{2}\sigma^2(\epsilon)(\pi^*_{W,\epsilon})^2 - c^*_{W,\epsilon}) \\ + \int_{\mathbb{R}\setminus\{0\}} \big([1 + \pi^*_{W,\epsilon}(e^z - 1)]^\gamma - 1 - \gamma\pi^*_{W,\epsilon}(e^z - 1)1_{|z|<1}\big)\, \nu_{N,\epsilon}(dz)]\Big\} dt$$

with

$$c^*_{W,\epsilon} = \frac{1-\gamma}{\delta - k_{W,\epsilon}(\gamma)}.$$

A Brownian approximation as in Section 2.3.2 results in an additional $\sigma^2(\epsilon)$-term also in $k_{W,\epsilon}(\gamma)$:

$$k_{W,\epsilon}(\gamma) = \gamma(r + (\hat{\mu}_{W,\epsilon} - r)\pi^*_{W,\epsilon}) - \frac{1}{2}\sigma^2(\epsilon)(\pi^*_{W,\epsilon})^2\gamma(1-\gamma) \\ + \int_{\mathbb{R}\setminus\{0\}} \big([1 + \pi^*_{W,\epsilon}(e^z - 1)]^\gamma - 1 - \gamma\pi^*_{W,\epsilon}(e^z - 1)\big)\nu_{N,\epsilon}(dz).$$

The convergence of the corresponding value function $V_{W,\epsilon}$ follows as for the case $V_{N,\epsilon}$ above. □

2.6 Convergence rate for the wealth process

Still remaining is the convergence of the wealth processes. Convergence in probability of the wealth processes is clear. For convergence in L_2 we can derive a rate which is, not surprisingly, proportional to $\sigma^2(\epsilon)$.

Proposition 2.6.1. *The wealth process $X_{N,\epsilon}$ converges to the original process X in L_2. For every $T < \infty$ it holds furthermore for the case of no consumption*

$$\sup_{t\in[0,T]} \mathbb{E}[|X(t) - X_{N,\epsilon}(t)|^2] \le K\sigma^2(\epsilon)$$

where K depends on T.

Proof. We can write the difference between the wealth processes as

$$X(t) - X_{N,\epsilon}(t) = \int_0^t (r + (\hat{\mu} - r)\pi)X(s) - (r + (\hat{\mu}_{N,\epsilon} - r)\pi_{N,\epsilon})X_{N,\epsilon}(s)ds \\ + \int_0^t cX(s) - c_{N,\epsilon}X_{N,\epsilon}(s)ds$$

$$
\begin{aligned}
&+\int_0^t\int_{\mathbb{R}\setminus\{0\}}\big(\pi X(s-)-\pi_{N,\epsilon}X_{N,\epsilon}(s-)1_{|z|>\epsilon}\big)(e^z-1)\tilde{N}(ds,dz)\\
=&\int_0^t(r+(\hat{\mu}-r)\pi)(X(s)-X_{N,\epsilon}(s))ds\\
&+\int_0^t X_{N,\epsilon}(s)\big((r+(\hat{\mu}-r)\pi)-(r+(\hat{\mu}_{N,\epsilon}-r)\pi_{N,\epsilon})\big)ds\\
&-\int_0^t c\big(X(s)-X_{N,\epsilon}(s)\big)ds-\int_0^t(c-c_{N,\epsilon})X_{N,\epsilon}(s)ds\\
&+\int_0^t\pi\big(X(s-)-X_{N,\epsilon}(s-)\big)\int_{\mathbb{R}\setminus\{0\}}(e^z-1)\tilde{N}(ds,dz)\\
&+\int_0^t\int_{\mathbb{R}\setminus\{0\}}X_{N,\epsilon}(s-)(\pi-\pi_{N,\epsilon}1_{|z|>\epsilon})(e^z-1)\tilde{N}(ds,dz).
\end{aligned}
$$

Then, it follows:

$$
\begin{aligned}
&\mathbb{E}[|X(t)-X_{N,\epsilon}(t)|^2]\\
&\quad\le c_1\mathbb{E}\Big[\Big(\int_0^t X(s)-X_{N,\epsilon}(s)ds\Big)^2\Big]\\
&\qquad+c_2\mathbb{E}\Big[\Big(\int_0^t X_{N,\epsilon}(s)ds\Big)^2\Big]\Big\{\big((\hat{\mu}-r)\pi-(\hat{\mu}_{N,\epsilon}-r)\pi_{N,\epsilon}\big)^2\Big\}\\
&\qquad+c_3\mathbb{E}\Big[\Big(\int_0^t\int_{\mathbb{R}\setminus\{0\}}(X(s-)-X_{N,\epsilon}(s-))(e^z-1)\tilde{N}(ds,dz)\Big)^2\Big]\\
&\qquad+c_4\mathbb{E}\Big[\Big(\int_0^t\int_{\mathbb{R}\setminus\{0\}}X_{N,\epsilon}(s-)(\pi-\pi_{N,\epsilon}1_{|z|>\epsilon})(e^z-1)\tilde{N}(ds,dz)\Big)^2\Big]
\end{aligned}
$$

for constants $c_1,...,c_4$. The constants denoted by c_1, c_2 are going to vary from step to step during this proof. For $t\le T$ we find by Cauchy-Schwarz

$$
\begin{aligned}
\mathbb{E}\Big[\Big(\int_0^t X(s)-X_{N,\epsilon}(s)ds\Big)^2\Big]&\le T\int_0^t\mathbb{E}\Big[\big(X(s)-X_{N,\epsilon}(s)\big)^2\Big]ds\\
\mathbb{E}\Big[\Big(\int_0^t X_{N,\epsilon}(s)ds\Big)^2\Big]&\le T\int_0^t\mathbb{E}\Big[X_{N,\epsilon}^2\Big]ds.
\end{aligned}
$$

Furthermore, we find

$$
\begin{aligned}
&\mathbb{E}\Big[\Big(\int_0^t\int_{\mathbb{R}\setminus\{0\}}(\pi(X(s-)-X_{N,\epsilon}(s-))(e^z-1)\nu(dz)ds)\Big)^2\Big]\\
&\quad=\mathbb{E}\Big[\int_0^t\int_{\mathbb{R}\setminus\{0\}}\big(X(s-)-X_{N,\epsilon}(s-)\big)^2(e^z-1)\nu(dz)ds\Big]\\
&\quad=\int_{\mathbb{R}\setminus\{0\}}(e^z-1)^2\nu(dz)\int_0^t\mathbb{E}\Big[(X(s)-X_{N,\epsilon}(s))^2\Big]ds, \qquad (2.6.1)
\end{aligned}
$$

$$
\mathbb{E}\Big[\Big(\int_0^t\int_{\mathbb{R}\setminus\{0\}}X_{N,\epsilon}(s-)(\pi-\pi_{N,\epsilon})1_{|z|>\epsilon}(e^z-1)\tilde{N}(ds,dz)\Big)^2\Big]
$$

$$
\begin{aligned}
&= \mathbb{E}\Big[\int_0^t \int_{\mathbb{R}\setminus\{0\}} X^2_{N,\epsilon}(s-)(\pi - \pi_{N,\epsilon} 1_{|z|>\epsilon})^2 (e^z - 1)^2 \nu(dz) ds\Big] \\
&= \int_0^t \mathbb{E}\Big[X^2_{N,\epsilon}(s)\Big] ds \int_{\mathbb{R}\setminus\{0\}} (\pi - \pi_{N,\epsilon} 1_{|z|>\epsilon})^2 (e^z - 1)^2 \nu(dz). \qquad (2.6.2)
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\mathbb{E}[|X(t) - X_{N,\epsilon}(t)|^2] \le c_1 &\int_0^t \mathbb{E}\Big[\big(X(s) - X_{N,\epsilon}(s)\big)^2\Big] ds \\
&+ c_2 \int_0^t \mathbb{E}\big[X^2_{N,\epsilon}(s)\big] ds \Big\{ \big((\hat{\mu} - r)\pi - (\hat{\mu}_{N,\epsilon} - r)\pi_{N,\epsilon}\big)^2 \\
&+ \int_{\mathbb{R}\setminus\{0\}} (\pi - \pi_{N,\epsilon} 1_{|z|>\epsilon})^2 (e^z - 1)^2 \nu(dz) \Big\}
\end{aligned}
$$

Also it follows that

$$
\begin{aligned}
&\int_{\mathbb{R}\setminus\{0\}} (\pi - \pi_{N,\epsilon} 1_{|z|>\epsilon})^2 (e^z - 1)^2 \nu(dz) \qquad (2.6.3) \\
&\quad = \int_{\mathbb{R}\setminus\{0\}} (\pi - \pi_{N,\epsilon} + \pi_{N,\epsilon} 1_{|z|<\epsilon})^2 (e^z - 1)^2 \nu(dz)\} \\
&\quad \le c_1 \int_{\mathbb{R}\setminus\{0\}} (e^z - 1)^2 \nu(dz) (\pi - \pi_{N,\epsilon})^2 + c_2 \int_{|z|<\epsilon} (e^z - 1)^2 \nu(dz) \\
&\quad \le c_1 |\pi - \pi_{N,\epsilon}|^2 + c_2 \sigma^2(\epsilon) \qquad (2.6.4)
\end{aligned}
$$

and

$$
\begin{aligned}
\big((\hat{\mu} - r)\pi - (\hat{\mu}_{N,\epsilon} - r)\pi_{N,\epsilon}\big)^2 &= \big((\hat{\mu} - r)(\pi - \pi_{N,\epsilon}) + \pi_{N,\epsilon}(\hat{\mu} - \hat{\mu}_{N,\epsilon})\big)^2 \\
&\le c_1 |\pi - \pi_{N,\epsilon}|^2 + c_2 |\hat{\mu} - \hat{\mu}_{N,\epsilon}|^2. \qquad (2.6.5)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\mathbb{E}[|X(t) - X_{N,\epsilon}(t)|^2] \le \tilde{c}_1 &\int_0^t \mathbb{E}[\big(X(s) - X_{N,\epsilon}(s)\big)^2] ds \\
&+ \tilde{c}_2 \big((|\pi - \pi_{N,\epsilon}|^2 + \sigma^2(\epsilon)\big) \int_0^t \mathbb{E}[X^2_{N,\epsilon}(s)] ds.
\end{aligned}
$$

Then it follows by Gronwall's inequality:

$$
\begin{aligned}
\mathbb{E}[|X(t) - X_{N,\epsilon}(t)|^2] &\le \tilde{c}_2 \int_0^t e^{\tilde{c}_1 (t-s)} \mathbb{E}[X^2_{N\epsilon}] ds \big(\sigma^2(\epsilon) + (\pi - \pi_{N\epsilon})^2\big) \\
&\le c_1 \sigma^2(\epsilon) + c_2 \sigma^4(\epsilon) \\
&\le c_1 \sigma^2(\epsilon).
\end{aligned}
$$

Hence, for every $T < \infty$ we conclude

$$
\sup_{t \in [0,T]} \mathbb{E}[|X(t) - X_{N,\epsilon}(t)|^2] \le K \sigma^2(\epsilon)
$$

where K depends on T, and the Proposition follows.

□

The convergence and convergence rates in this paper are analysed for the specific case of power utility in a Merton framework. The proofs, especially for the convergence rate of the optimal control, depend on features of the concrete form of the solutions in this specific setting. In a more general setting a concrete solution is not available. Additionally it is not clear if the optimal control and the consumption rate are constant in time.

CHAPTER 3

ARTICLE 2: "ON STABILITY TO MODEL RISK OF OPTIONS IN A BIVARIATE LÉVY MARKET"

Fred Espen Benth, Giulia Di Nunno, Asma Khedher and Maren Diane Schmeck

Abstract

We consider options that depend on multiple assets, in particular spread options. We assume that the possibly dependent assets are driven by a Lévy process with infinite activity. Then we substitute the small jumps by a scaled Brownian motion and explore the impact on the corresponding option price by quantifying an error estimate. Hereby, we extract the contribution from the dependency of the assets. If one approximates under the real world measure, the risk neutral pricing measure will be model dependent, and contributes additionally to the convergence rate.[1]

3.1 Introduction

There are many options that depend on several underlying assets. For instance, spread options are written on the difference of two underlying assets $S^{(2)}(t) - S^{(1)}(t)$, $t \geq 0$. The asset price dynamics often includes jumps, which can be driven by a Lévy process with infinite activity. Asmussen and Rosinski [5] prove that the small jumps have a central limit type behavior towards a Brownian motion, which was extended to the multivariate case by Cohen and Rosinski [30]. Thus it is not a simple task to decide how the small movements of the assets price dynamics behave, as a continuous process or by infinitely many jumps. Consider an investor's view, who wants to choose a model for the (multidimensional) asset price dynamics for the purpose of option pricing. She is therefore interested in the effect of the choice of the model in the option price and in a quantification of an upper bound for the difference in the two prices. Benth, Di Nunno and Khedher [9, 10, 11] address this problem

[1] This article is a completely revised version of Benth, Di Nunno, Khedher and Schmeck [12].

in the one-dimensional case, which is extended by Khedher [44] to a multidimensional setting with independent assets. Here, we include dependency into the discussion, as the price of an option on several underlying might crucially depend on their correlation. For example, the spread of two options at a future time T depends on whether the asset prices move into the same direction or not. We specify the influence of the dependency as a part of the convergence rate. Our results are valid for options that admit a Fourier transform based pricing formula as in Eberlein, Glau and Papapantoleon [34].

An investor specifies her model under the real world measure $\mathbb{P}$ and would therefore also approximate under $\mathbb{P}$. This results in a risk neutral pricing measure $\mathbb{Q}$ that is dependent on the choice of model, and the price difference will become dependent on the characteristics of the measure change, as the parameter of the Esscher transform. Benth, Di Nunno and Khedher [10] show robustness of the option prices under different types of measure changes. Approximating directly under $\mathbb{Q}$ is useful for simulation purposes, see Cont and Tankov [31].

We anlayse the stability of an option with a bivariate underlying when approximating under $\mathbb{Q}$. Afterwards, we examine the case of spread options, and approximate the small jumps before and after a measure change. Therefore, we state a Margrabe type formula for spread options, see Margrabe [48] and Carmona and Durrleman [28] for spread options in continuous models. The formula allows to move from pricing the spread option written on a bivariate process to pricing a European option written on a one-dimensional process and is based on an appropriate change of measure. Margrabe's formula received lately a lot of attention in academic literature. Cheang and Chiarella [29] examine it in a jump diffusion setting and Eberlein, Papapantoleon and Shiryaev [36] study the problem of valuation of options depending on several assets using a duality approach. This includes a formula for the valuation of spread option written on exponential semimartingales in terms of the triplet of predictable characteristics of a one-dimensional semimartingale under the dual measure. Another approach to price spread options can be found in Borovkova, Permana and v.d. Weide [25].

Benth, Di Nunno and Khedher [9, 10, 11] and Khedher [44] find the same rates for the case of approximation as when one simply neglects small jumps. Here we specify their results and get a different rate if we approximate them by a Brownian motion.

The paper is organised as follows. In Section 3.2 we study the Gaussian approximation of the small jumps in a bivariate setting and prove convergence in $\mathcal{L}^1$ of the underlying Lévy processes. Then we give convergence rates for the option price on a bivariate underlying based on Fourier transformation. We move on to pricing and stability of spread options. In Section 3.3 we state Margrabe's formula in the jump-diffusion setting and derive convergence rates if one approximates under $\mathcal{P}$ or under Margrabe's pricing measure.

3.2 Brownian approximation of a bivariate Lévy process

Before we go on to option pricing, we discuss the Brownian approximation of the small jumps in a bivariate setting and include dependency between the two processes. We consider a bivariate Lévy process $L = (L^{(1)}, L^{(2)})$ with Lévy measure $\nu(dz) = \nu(dz_1, dz_2)$, where the components have the Lévy-Itô decomposition

$$L^{(1)}(t) = a_1 t + \int_0^t \int_{|z|<1} z_1 \tilde{N}(ds, dz_1, dz_2) + \int_0^t \int_{|z|\geq 1} z_1 N(ds, dz_1, dz_2) \qquad (3.2.1)$$

$$L^{(2)}(t) = a_2 t + \int_0^t \int_{|z|<1} z_2 \tilde{N}(ds, dz_1, dz_2) + \int_0^t \int_{|z|\geq 1} z_2 N(ds, dz_1, dz_2) \ ,$$

with $|z| = \sqrt{z_1^2 + z_2^2}$. Furthermore, let $B = (B^{(1)}, B^{(2)})$ be a bivariate standard Brownian motion, where $B^{(1)}$ and $B^{(2)}$ are independent and let $\alpha(\epsilon) \in \mathbb{R}^{2\times 2}$ be a scaling matrix depending on ϵ with $0 < \epsilon < 1$. Neglecting in L the jumps with absolute size smaller than ϵ and approximating them by $\alpha(\epsilon)B(t)$ leads to an approximation L_ϵ of L. Here

$$\alpha(\epsilon) = \begin{pmatrix} \alpha_1(\epsilon) & \alpha_2(\epsilon) \\ \alpha_2(\epsilon) & \alpha_3(\epsilon) \end{pmatrix}$$

and the components of L_ϵ are given by

$$\begin{aligned} L_\epsilon^{(1)}(t) &= a_1 t + \alpha_1(\epsilon)B^{(1)}(t) + \alpha_2(\epsilon)B^{(2)}(t) \\ &\quad + \int_0^t \int_{\epsilon\leq|z|<1} z_1 \tilde{N}(ds, dz_1, dz_2) + \int_0^t \int_{|z|\geq 1} z_1 N(ds, dz_1, dz_2) \\ L_\epsilon^{(2)}(t) &= a_2 t + \alpha_2(\epsilon)B^{(1)}(t) + \alpha_3(\epsilon)B^{(2)}(t) \\ &\quad + \int_0^t \int_{\epsilon\leq|z|<1} z_2 \tilde{N}(ds, dz_1, dz_2) + \int_0^t \int_{|z|\geq 1} z_2 N(ds, dz_1, dz_2). \end{aligned} \tag{3.2.2}$$

Denote by ν_ϵ the Lévy measure of L_ϵ, where

$$\nu_\epsilon(dz_1, dz_2) := \begin{cases} \nu(dz_1, dz_2)\,, & |z| \geq \epsilon \\ 0\,, & \text{otherwise.} \end{cases}$$

Define

$$\begin{aligned} \sigma_1^2(\epsilon) &= \int_{|z|<\epsilon} z_1^2 \nu(dz_1, dz_2) \\ \sigma_2^2(\epsilon) &= \int_{|z|<\epsilon} z_2^2 \nu(dz_1, dz_2) \\ \sigma_{12}(\epsilon) &= \int_{|z|<\epsilon} z_1 z_2 \nu(dz_1, dz_2) \end{aligned} \tag{3.2.3}$$

and denote

$$\Sigma(\epsilon) = \begin{pmatrix} \sigma_1^2(\epsilon) & \sigma_{12}(\epsilon) \\ \sigma_{12}(\epsilon) & \sigma_2^2(\epsilon) \end{pmatrix} \ .$$

Choose the scaling coefficients of the Brownian motion such that

$$\alpha(\epsilon)\alpha(\epsilon)^T = \Sigma(\epsilon) \ . \tag{3.2.4}$$

This ensures that we keep the variance-covariance matrix of the original process, as (3.2.4) is equivalent to

$$Var(\alpha(\epsilon)B(1)) = \Sigma(\epsilon) \ .$$

Written component wise, we get following set of equations for the coefficients $\alpha_1(\epsilon)$, $\alpha_2(\epsilon)$, and $\alpha_3(\epsilon)$

$$\begin{aligned}\alpha_1^2(\epsilon)+\alpha_2^2(\epsilon)&=\sigma_1^2(\epsilon)\\ \alpha_1(\epsilon)\alpha_2(\epsilon)+\alpha_2(\epsilon)\alpha_3(\epsilon)&=\sigma_{12}(\epsilon)\\ \alpha_2^2(\epsilon)+\alpha_3^2(\epsilon)&=\sigma_2^2(\epsilon).\end{aligned}\tag{3.2.5}$$

We know that $\sigma_1^2(\epsilon)$ and $\sigma_2^2(\epsilon)$ vanish when ϵ goes to 0. Therefore $\alpha_1(\epsilon)$, $\alpha_2(\epsilon)$, and $\alpha_3(\epsilon)$ converge also to 0 when ϵ goes to 0. We use this to prove the following convergence result.

Proposition 3.2.1. *Let the process* $(L^{(1)},L^{(2)})$ *and* $(L_\epsilon^{(1)},L_\epsilon^{(2)})$ *be defined as in* (3.2.1) *and* (3.2.2). *Then, for every* $t\geq 0$,

$$\lim_{\epsilon\to 0}(L_\epsilon^{(1)}(t),L_\epsilon^{(2)}(t))=(L^{(1)}(t),L^{(2)}(t))\quad \mathbb{P}-a.s.$$

The limit above also holds in $\mathcal{L}^1(\Omega,\mathcal{F},\mathbb{P})$ *with*

$$\mathbb{E}\left[|L_\epsilon^{(1)}(t)-L^{(1)}(t)|\right]\leq\big(\alpha_1(\epsilon)+\alpha_2(\epsilon)+\sigma_1(\epsilon)\big)\sqrt{t}$$

and

$$\mathbb{E}\left[|L_\epsilon^{(2)}(t)-L^{(2)}(t)|\right]\leq\big(\alpha_2(\epsilon)+\alpha_3(\epsilon)+\sigma_2(\epsilon)\big)\sqrt{t}\,.$$

Proof. The $\mathbb{P}$-a.s. convergence follows directly from the Lévy-Khintchine formula (see Thm. 19.2 in Sato [53]). Concerning the $\mathcal{L}^1$-convergence, we argue as follows. The combined application of the triangle and Cauchy-Schwarz inequalities gives

$$\begin{aligned}&\mathbb{E}\left[|L^{(1)}(t)-L_\epsilon^{(1)}(t)|\right]\\ &\quad=\mathbb{E}\left[\left|\alpha_1(\epsilon)B^{(1)}(t)+\alpha_2(\epsilon)B^{(2)}(t)-\int_0^t\int_{|z|<\epsilon}z_1\,\widetilde{N}(ds,dz_1,dz_2)\right|\right]\\ &\quad\leq\alpha_1(\epsilon)\mathbb{E}\left[\left|B^{(1)}(t)\right|\right]+\alpha_2(\epsilon)\mathbb{E}\left[\left|B^{(2)}(t)\right|\right]\\ &\qquad\quad+\mathbb{E}\left[\left|\int_0^t\int_{|z|<\epsilon}z_1\,\widetilde{N}(ds,dz_1,dz_2)\right|\right]\\ &\quad\leq\alpha_1(\epsilon)\mathbb{E}\left[\left|(B^{(1)}(t))^2\right|\right]^{\frac12}+\alpha_2(\epsilon)\mathbb{E}\left[\left|(B^{(2)}(t))^2\right|\right]^{\frac12}\\ &\qquad\quad+\mathbb{E}\left[\left|\int_0^t\int_{|z|<\epsilon}z_1\,\widetilde{N}(ds,dz_1,dz_2)\right|^2\right]^{\frac12}\\ &\quad\leq\big(\alpha_1(\epsilon)+\alpha_2(\epsilon)+\sigma_1(\epsilon)\big)\sqrt{t}.\end{aligned}$$

The coefficients σ_1, σ_2, and σ_{12} converge to 0 when ϵ goes to 0. Therefore, from (3.2.5), we deduce that the coefficients $\alpha_1(\epsilon)$, $\alpha_2(\epsilon)$, and $\alpha_3(\epsilon)$ go to 0 when ϵ goes to 0. In the same manner, we can prove that $\mathbb{E}[|L^{(2)}-L_\epsilon^{(2)}|]\leq\Big(\alpha_2(\epsilon)+\alpha_3(\epsilon)+\sigma_2(\epsilon)\Big)\sqrt{t}$ and the result follows. □

In the proof of the Proposition 3.2.1 it is enough that the coefficients of the matrix $\alpha(\epsilon)$ converge to 0 for $\epsilon\to 0$. That is, as long as $\alpha(\epsilon)$ converges to 0, we could have chosen the scaling matrix differently.

3.2.1 Stability of option prices with a bivariate underlying

We are interested in the price of an option with exercise time T of the form

$$C_X = \mathbb{E}[f(X(T))] \,, \tag{3.2.6}$$

where X is a jump-diffusion in $\mathbb{R}^2$ and $f : \mathbb{R}^2 \to \mathbb{R}_+$ a payoff function. Eberlein, Glau and Papapantoleon [34] establish valuation formulas for options using the Fourier transform $\widehat{f}$ of the payoff functions. We denote by $P_{X(T)}(dx)$ the probability of $X(T)$ and let $R \in \mathbb{R}^2$ be a damping factor. Assume the following conditions are satisfied

1. $\mathrm{e}^{-Rx} f(x) \in \mathcal{L}^1(\mathbb{R}^2)$,
2. $\widehat{\mathrm{e}^{-Rx} f(x)} \in \mathcal{L}^1(\mathbb{R}^2)$,
3. $\mathrm{e}^{Rx} P_{X(T)}(dx) \in \mathcal{L}^1(\mathbb{R}^2)$.

Then the price C_X of the option (3.2.6) at time 0 is given by

$$C_X = \frac{e^{-R.s}}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-iu.s} \phi_{X_T}(u - iR) \widehat{f}(iR - u) du \tag{3.2.7}$$

with $s = (s_1, s_2) \in \mathbb{R}^2$ and $s^i = -logS_0^{(i)}$ (see Theorem 3.2 in Eberlein, Glau and Papapantoleon [34]).

We now want to examine the difference between the option prices

$$C_L = \mathbb{E}[f(L(T))] \qquad \text{and} \qquad C_{L_\epsilon} = \mathbb{E}[f(L_\epsilon(T))] \,.$$

Using (3.2.7), it is

$$|C_L - C_{L_\epsilon}| = \left| \frac{e^{-R.s}}{(2\pi)^2} \int_{\mathbb{R}^2} e^{-iu.s} \Big(\phi_{L(T)}(u - iR) - \phi_{L_\epsilon(T)}(u - iR) \Big) \widehat{f}(iR - u) du \right| \,.$$

Therefore we need to quantify the difference between the corresponding characteristic functions, which are given by

$$\phi_{L(t)}(u) = \exp\Big\{ t(iu.a + \int_{\mathbb{R}^2} \mathrm{e}^{iu.z} - 1 - iu.z 1_{|z|<1} \nu(dz)) \Big\},$$
$$\phi_{L_\epsilon(t)}(u) = \exp\Big\{ t(iu.a - \frac{1}{2} u.\Sigma(\epsilon) u + \int_{|z| \geq \epsilon} e^{iu.z} - 1 - iu.z 1_{|z|<1} \nu(dz) \Big\} \,.$$

Proposition 3.2.2. *We have*

$$|\phi_{L(t)}(u) - \phi_{L_\epsilon(t)}(u)| \leq c_1(u) \epsilon \sigma_1^2(\epsilon) + c_2(u) \epsilon \sigma_2^2(\epsilon) + c_{12}(u) \epsilon \sigma_{12}(\epsilon)$$

with functions $c_1(u), c_2(u), c_{12}(u)$ *depending on* u.

Proof. Define a function f as the difference of the log-characteristic functions of $L_\epsilon(1)$ and $L(1)$

$$f(u, \epsilon) \;\; := \;\; -\frac{1}{2} u.\Sigma(\varepsilon) u - \int_{|z|<\epsilon} \Big(e^{iu.z} - 1 - iu.z 1_{|z|<1} \Big) \nu(dz) \tag{3.2.8}$$

$$= \quad -\int_{|z|<\epsilon} \Big(e^{i(u_1z_1+u_2z_2)} - 1 - i(u_1z_1+u_2z_2) + \frac{1}{2}(u_1^2z_1^2 + u_2^2z_2^2 + 2u_1u_2z_1z_2) \Big) \nu(dz) \ .$$

Then

$$\begin{aligned} |\phi_{L(t)}(u) - \phi_{L_\epsilon(t)}(u)| &= \phi_{L(t)}(u)|1 - \exp\{tf(u,\epsilon)\}| \\ &\leq \phi_{L(t)}(u)t|f(u,\epsilon)|\exp\{t|f(u,1)|\} \ . \end{aligned}$$

With Taylor expansions we find that

$$\begin{aligned} e^{i(u_1z_1+u_2z_2)} \quad = \quad & 1 + i(u_1z_1+u_2z_2) - \frac{1}{2}(u_1^2z_1^2 + u_2^2z_2^2 + 2u_1u_2z_1z_2) \\ & + \sum_{n_2=3}^{\infty} \frac{(iu_2z_2)^{n_2}}{n_2!} + \sum_{n_2=2}^{\infty} \frac{(iu_1z_1)}{1!}\frac{(iu_2z_2)^{n_2}}{n_2!} \\ & + \sum_{n_2=1}^{\infty} \frac{(iu_1z_1)^2}{2!}\frac{(iu_2z_2)^{n_2}}{n_2!} + \sum_{n_1=3}^{\infty}\sum_{n_2=0}^{\infty} \frac{(iu_1z_1)^{n_1}}{n_1!}\frac{(iu_2z_2)^{n_2}}{n_2!} \ . \end{aligned} \tag{3.2.9}$$

For $|z| < \epsilon$ we have

$$\Big| \sum_{n_2=3}^{\infty} \frac{(iu_2z_2)^{n_2}}{n_2!} \Big| = \Big| (iu_2z_2)^3 \sum_{n_2=0}^{\infty} \frac{(iu_2z_2)^{n_2}}{(n_2+3)!} \Big| \leq |u_2z_2|^3 e^{|u_2|} \ .$$

We estimate the other summand in (3.2.9) analogously and it follows

$$\begin{aligned} |f(u,\epsilon)| \quad \leq \quad & |u_2^3|e^{|u_2|} \int_{|z|<\epsilon} |z_2|^3 \nu(dz) + |u_1u_2^2|e^{|u_2|} \int_{|z|<\epsilon} |z_1z_2^2| \nu(dz) \\ & + |u_1^2u_2|e^{|u_2|} \int_{|z|<\epsilon} |z_1^2z_2| \nu(dz) + |u_1^3|e^{|u_1|+|u_2|} \int_{|z|<\epsilon} |z_1|^3 \nu(dz) \\ \leq \quad & c_1(u) \int_{|z|<\epsilon} |z_2|^3 \nu(dz) + c_2(u) \int_{|z|<\epsilon} |z_1z_2^2| \nu(dz) \\ & + c_3(u) \int_{|z|<\epsilon} |z_1^2z_2| \nu(dz) + c_4(u) \int_{|z|<\epsilon} |z_1|^3 \nu(dz) \ , \end{aligned}$$

for constants $c_i(u)$, $i = 1, ..., 4$ depending on u. Then the Proposition follows. □

Using Proposition 3.2.2 we find immediately

Proposition 3.2.3. *Assume that*

$$\begin{aligned} d_1 &= \frac{e^{-R.s}}{(2\pi)^2} \int_{\mathbb{R}^2} |\widehat{f}(iR-u)||c_1(u-iR)| \, du \\ d_2 &= \frac{e^{-R.s}}{(2\pi)^2} \int_{\mathbb{R}^2} |\widehat{f}(iR-u)||c_2(u-iR)| \, du \\ d_{12} &= \frac{e^{-R.s}}{(2\pi)^2} \int_{\mathbb{R}^2} |\widehat{f}(iR-u)||c_{12}(u-iR)| \, du \end{aligned}$$

are finite. Then

$$|C_L - C_{L_\epsilon}| \leq d_1 \epsilon \sigma_1^2(\epsilon) + d_2 \epsilon \sigma_2^2(\epsilon) + d_{12} \epsilon \sigma_{12}(\epsilon) \ .$$

The covariance $\sigma_{12}(\epsilon)$ is now also included in the convergence rate, additionally to the variance of $L^{(1)}$ and $L^{(2)}$.

The approximation (3.2.2) is reached in two steps, truncating the small jumps and approximating them by a Brownian motion. Then in (3.2.8) the terms of the Taylor approximation of (3.2.9) cancel away up till order 2. In fact, we have chosen the scaling coefficients of the Brownian motion exactly such that the second order terms disappear. Without the Brownian approximation the second order terms do not cancel. So consider the following approximation of L in (3.2.1) where one simply neglects the small jumps

$$L_{N,\epsilon}(t) = at + \int_0^t \int_{\epsilon \leq |z| < 1} z\tilde{N}(ds, dz) + \int_0^t \int_{|z| \geq 1} zN(ds, dz)$$

with characteristic function

$$\phi_{L_{N,\epsilon}(t)}(u) = \exp\left\{ iu.at + t \int_{|z| \geq \epsilon} \left(e^{iu.z} - 1 - iu.z1_{|z|<1} \right) \nu(dz) \right\} \ .$$

The difference between the log-characteristic functions of $L_{N,\epsilon}(1)$ and $L(1)$ is given by

$$f(u, \epsilon) = - \int_{|z| < \epsilon} \left(e^{iu.z} - 1 - iu.z1_{|z|<1} \right) \nu(dz)$$

Using that

$$\begin{aligned} e^{i(u_1 z_1 + u_2 z_2)} &= 1 + i(u_1 z_1 + u_2 z_2) + (iu_2 z_2)^2 \sum_{n_2=0}^{\infty} \frac{(iu_2 z_2)^{n_2}}{(n_2 + 2)!} \\ &\quad + (iu_1 z_1)(iu_2 z_2) \sum_{n_2=0}^{\infty} \frac{(iu_2 z_2)^{n_2}}{(n_2 + 1)!} + (iu_1 z_1)^2 \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} \frac{(iu_1 z_1)^{n_1}}{(n_1 + 2)!} \frac{(iu_2 z_2)^{n_2}}{n_2!} \ . \end{aligned}$$

we find analogously to the proof of Proposition 3.2.2

$$\begin{aligned} |f(u, \epsilon)| &\leq |u_2^2| e^{|u_2|} \int_{|z| < \epsilon} |z_2|^2 \nu(dz) + |u_1 u_2^2| e^{|u_2|} \int_{|z| < \epsilon} |z_1 z_2| \nu(dz) \\ &\quad + |u_1^2| e^{|u_1| + |u_2|} \int_{|z| < \epsilon} |z_1|^2 \nu(dz) \ . \end{aligned}$$

Therefore it follows

$$|\phi_{L(t)}(u) - \phi_{L_{N,\epsilon}(t)}(u)| \leq c_1(u) \sigma_1^2(\epsilon) + c_2(u) \sigma_2^2(\epsilon) + c_{12}(u) \sigma_{12}(\epsilon)$$

for different c_1, c_2 and c_{12} depending on u.

3.3 Pricing and stability of spread options

A spread is defined by the difference of the two underlying asset prices $S^{(2)}(t) - S^{(1)}(t)$, $t \geq 0$. Thus, the payout function of a European spread option with strike 0 at maturity date T is given by

$$\max(S^{(2)}(T) - S^{(1)}(T), 0) \; . \tag{3.3.1}$$

In the following we state a Margrabe type formula for a spread option written on a bivariate jump-diffusion (see Section 5.2 in Carmona and Durrleman [28] for spread options written on continuous processes). We consider a spread option written on

$$S^{(1)}(t) = S^{(1)}(0)\mathrm{e}^{L^{(1)}(t)} \qquad \text{and} \qquad S^{(2)}(t) = S^{(2)}(0)\mathrm{e}^{L^{(2)}(t)} \; , \tag{3.3.2}$$

where $L = (L^{(1)}, L^{(2)})$ is a bivariate Lévy process with characteristic triplet (a, Σ, ν) under the real world measure $\mathbb{P}$.

Our computations will be based on the Esscher transform of Gerber and Shiu [41] for options on several risky assets. The Esscher probability $\mathbb{Q}_\theta$ is defined by means of the Esscher transform as follows (see Gerber and Shiu [41])

$$\frac{d\mathbb{Q}_\theta}{d\mathbb{P}}\Big|_{\mathcal{F}_t} = \frac{\mathrm{e}^{\theta . L(t)}}{M_t(\theta)} \; , \tag{3.3.3}$$

where the moment generating function of L is given by

$$M_t(\theta) = \mathbb{E}_{\mathbb{P}}[\mathrm{e}^{\theta . L(t)}] \; .$$

The transform depends on the parameter $\theta \in \mathbb{R}^2$. By Theorems 33.1 and 33.2 in Sato [53], the process L is again a Lévy process under $\mathbb{Q}_\theta$ with characteristic triple $(a_\theta, \Sigma, \nu_\theta)$

$$\nu_\theta(dz) = \mathrm{e}^{\theta . z}\nu(dz)$$
$$a_\theta = a + \Sigma.\theta + \int_{|z|<1} z(e^{\theta z} - 1)\nu(dz)$$

In order for (3.3.3) to be well-defined, we must impose exponential integrability conditions on $L(1)$. Therefore suppose that there exists a constant $c > 0$ such that

$$\int_{\mathbb{R}^2} \mathrm{e}^{c.z}\, \nu(dz) < \infty \, , \tag{3.3.4}$$

for all $|z| \leq c$. This ensures finite exponential moments for $L(1)$ up to order c.

One needs to define a risk neutral Esscher probability. Therefore the parameter θ is determined such that the discounted price process $\mathrm{e}^{-rt}S^{(i)}(t)$ is a martingale for $i = 1, 2$. Hence

$$S^{(i)}(0) = \mathbb{E}_{\mathbb{Q}_\theta}[\mathrm{e}^{-rt}S^{(i)}(t)] \; , \tag{3.3.5}$$

which is equivalent to

$$e^{rt} = \mathbb{E}_{\mathbb{Q}_\theta}[\mathrm{e}^{L^{(i)}(t)}] = \mathbb{E}_{\mathbb{P}}\Big[\frac{\mathrm{e}^{L^{(i)}(t)+\theta . L(t)}}{M_t(\theta)}\Big] = \frac{M_t(\mathbf{1}_i + \theta)}{M_t(\theta)}, \tag{3.3.6}$$

where $\mathbf{1}_i$ denotes the ith unit vector. The existence and uniqueness of the parameter $\theta = (\theta_1, \theta_2)$, which verifies (3.3.6), is proven in Gerber and Shiu [41]. By the risk neutral valuation rule, the price of the spread option is then given by

$$C = \mathrm{e}^{-rT}\mathbb{E}_{\mathbb{Q}_\theta}\Big[\max\big(S^{(2)}(T) - S^{(1)}(T), 0\big)\Big] \,.$$

Define a new pricing measure $\mathbb{Q}_{\theta+\mathbf{1}_1}$ as an Esscher transformation with parameter $\mathbf{1}_1 = (1,0)$ with respect to $\mathbb{Q}_\theta$:

$$\frac{d\mathbb{Q}_{\theta+\mathbf{1}_1}}{d\mathbb{Q}_\theta}\bigg|_{\mathcal{F}_t} = \frac{\mathrm{e}^{\mathbf{1}_1.L(t)}}{\mathbb{E}_{\mathbb{Q}_\theta}[\mathrm{e}^{\mathbf{1}_1.L(t)}]} = e^{-rt+L^{(1)}(t)} \,. \tag{3.3.7}$$

The last equality follows from (3.3.5). Furthermore, it is

$$\frac{d\mathbb{Q}_{\theta+\mathbf{1}_1}}{d\mathbb{Q}_\theta}\frac{d\mathbb{Q}_\theta}{d\mathbb{P}}\bigg|_{\mathcal{F}_t} = e^{-rt+L^{(1)}(t)}\frac{\mathrm{e}^{\theta.L(t)}}{M_t(\theta)}\bigg|_{\mathcal{F}_t} = \frac{\mathrm{e}^{(\theta+\mathbf{1}_1).L(t)}}{M_t(\theta+\mathbf{1}_1)}\bigg|_{\mathcal{F}_t} = \frac{d\mathbb{Q}_{\theta+\mathbf{1}_1}}{d\mathbb{P}}\bigg|_{\mathcal{F}_t}, \tag{3.3.8}$$

using (3.3.6). Thus, $\mathbb{Q}_{\theta+\mathbf{1}_1}$ corresponds to a measure defined through an Esscher transform with respect to $\mathbb{P}$ with parameter $\theta + \mathbf{1}_1$.

We now give a version Margrabe's formula, that transforms the price of a spread option, written on a bivariate underlying, into the price of a European option on a one-dimensional underlying.

Proposition 3.3.1. *The price C of a spread option with strike $K = 0$ and maturity T is given by*

$$C = S^{(1)}(0)\mathbb{E}_{\mathbb{Q}_{\theta+\mathbf{1}_1}}\left[\max\left(\frac{S^{(2)}(T)}{S^{(1)}(T)} - 1, 0\right)\right] \,.$$

Proof. The price of a zero-exercise spread option is given by

$$\begin{aligned} C &= \mathrm{e}^{-rT}\mathbb{E}_{\mathbb{Q}_\theta}\left[\max\left(S^{(2)}(T) - S^{(1)}(T), 0\right)\right] \\ &= \mathrm{e}^{-rT}\mathbb{E}_{\mathbb{Q}_\theta}\left[\max\left(\frac{S^{(2)}(T)}{S^{(1)}(T)} - 1, 0\right)S^{(1)}(T)\right] \\ &= S^{(1)}(0)\mathbb{E}_{\mathbb{Q}_{\theta+\mathbf{1}_1}}\left[\max\left(\frac{S^{(2)}(T)}{S^{(1)}(T)} - 1, 0\right)\right] \,, \end{aligned}$$

where the last equation follows using the definition of $S^{(1)}$ in (3.3.2) and the density (3.3.7). □

With Itô calculus it follows that the process $\frac{S^{(2)}(t)}{S^{(1)}(t)}$ is again a geometric Lévy process under $\mathbb{Q}_{\theta+\mathbf{1}_1}$ and it is given by the following stochastic differential equation

$$\begin{aligned} d\left(\frac{S^{(2)}(t)}{S^{(1)}(t)}\right) = \frac{S^{(2)}(t)}{S^{(1)}(t)}\Big\{&(\sigma_{21} - \sigma_{11})dB^{(1)}_{\theta+\mathbf{1}_1}(t) + (\sigma_{22} - \sigma_{12})dB^{(2)}_{\theta+\mathbf{1}_1}(t) \\ &+ \int_{\mathbb{R}_0^2} e^{z_2 - z_1} - 1\widetilde{N}_{\theta+\mathbf{1}_1}(dt, dz_1, dz_2)\Big\}. \end{aligned}$$

The solution of the latter is given by

$$\frac{S^{(2)}(t)}{S^{(1)}(t)} = \frac{S^{(2)}(0)}{S^{(1)}(0)} \exp(X(t)) \,,$$

where

$$\begin{aligned} X(t) = & -\frac{1}{2}(\sigma_{21}-\sigma_{11})^2 - \frac{1}{2}(\sigma_{22}-\sigma_{12})^2 + (\sigma_{21}-\sigma_{11})B^{(1)}_{\theta+\mathbf{1}_1}(t) + (\sigma_{22}-\sigma_{12})B^{(2)}_{\theta+\mathbf{1}_1}(t) \\ & - t\int_{\mathbb{R}^2_0} \Big(e^{z_2-z_1} - 1 - (z_2-z_1)1_{|z|<1}\Big)\nu_{\theta+\mathbf{1}_1}(dz_1,dz_2) \\ & + t\int_{|z|<1}(z_2-z_1)\widetilde{N}_{\theta+\mathbf{1}_1}(ds,dz_1,dz_2) + t\int_{|z|\geq 1}(z_2-z_1)N(ds,dz_1,dz_2). \end{aligned} \tag{3.3.9}$$

Here, $B_{\theta+\mathbf{1}_1}(t)$ is a bivariate standard Brownian motion under $\mathbb{Q}_{\theta+\mathbf{1}_1}$ and $\widetilde{N}_{\theta+\mathbf{1}_1}(ds,dz_1,dz_2)$ is compensated with $\nu_{\theta+\mathbf{1}_1}(dz_1,dz_2)$.

3.3.1 Approximation under $\mathbb{Q}_{\theta+\mathbf{1}_1}$

Now, we consider the price of an spread option, where $S^{(1)}$ an $S^{(2)}$ are driven by a purely non-Gaussian Lévy process with infinite activity. Proposition 3.3.1 gives the price of this spread option as

$$C = S^{(1)}(0)\mathbb{E}_{\mathbb{Q}_{\theta+\mathbf{1}_1}}\left[\max\left(\frac{S^{(2)}(T)}{S^{(1)}(T)} - 1, 0\right)\right] \,,$$

where

$$\frac{S^{(2)}(t)}{S^{(1)}(t)} = \exp(X(t)),$$

$$X(t) := at + \int_0^t\int_{|z|<1}(z_2-z_1)\tilde{N}_{\theta+\mathbf{1}_1}(ds,dz_1,dz_2) + \int_0^t\int_{|z|\geq 1}(z_2-z_1)N(ds,dz_1,dz_2).$$

for some $a \in \mathbb{R}$. Now, approximate the small jumps of X by a scaled Brownian motion under the measure $\mathbb{Q}_{\theta+\mathbf{1}_1}$

$$\begin{aligned} X_\epsilon(t) \quad = \quad & at + \sigma_X(\epsilon)B_{\theta+\mathbf{1}_1}(t) + \int_0^t\int_{\epsilon\leq|z|<1}(z_2-z_1)\tilde{N}_{\theta+\mathbf{1}_1}(ds,dz_1,dz_2) \\ & + \int_0^t\int_{|z|\geq 1}(z_2-z_1)N(ds,dz_1,dz_2) \,, \end{aligned}$$

where

$$\sigma^2_X(\epsilon) = \int_{|z|<\epsilon}(z_2-z_1)^2\nu_{\theta+\mathbf{1}_1}(dz_1,dz_2) \,.$$

The approximated spread option price is then given by

$$C_\epsilon = S^{(1)}(0)\mathbb{E}_{\mathbb{Q}_{\theta+\mathbf{1}_1}}\left[\max\left(\frac{S^{(2)}_\epsilon(T)}{S^{(1)}_\epsilon(T)} - 1, 0\right)\right] \,,$$

where

$$\frac{S_\epsilon^{(2)}(t)}{S_\epsilon^{(1)}(t)} = \exp(X_\epsilon(t)) \ .$$

A straight forward calculation leads to the following characteristic functions of $X(t)$ and $X_\epsilon(t)$ under $\mathbb{Q}_{\theta+\mathbf{1}_1}$

$$\begin{aligned}
\phi_{X(t)}(u) &= \exp\left\{t\Big(iau + \int_{\mathbb{R}^2}\left\{e^{iu(z_2-z_1)} - 1 - iu(z_2-z_1)1_{|z|<1}\right\}\nu_{\theta+\mathbf{1}_1}(dz_1,dz_2)\Big)\right\}, \\
\phi_{X_\epsilon(t)}(u) &= \exp\left\{t\Big(iau - \frac{1}{2}u^2\sigma_X^2(\epsilon)\right. \\
&\qquad\left. + \int_{|z|\geq\epsilon}\left\{e^{iu(z_2-z_1)} - 1 - iu(z_2-z_1)1_{|z|<1}\right\}\nu_{\theta+\mathbf{1}_1}(dz_1,dz_2)\Big)\right\} \ .
\end{aligned}$$

We have then the following Proposition.

Proposition 3.3.2. *It holds that*

$$|\phi_{X(t)}(u) - \phi_{X_\epsilon(t)}(u)| \leq c(u) \int_{|z|<\epsilon} |z_2 - z_1|^3 \nu_{\theta+\mathbf{1}_1}(dz_1,dz_2) \ ,$$

where $c(u)$ depends on u.

Proof. Define the function

$$f(u,\epsilon) = -\frac{1}{2}\sigma_X^2(\epsilon)u^2 - \int_{|z|<\epsilon} e^{iu(z_2-z_1)} - 1 - iu(z_2-z_1)\nu_{\theta+\mathbf{1}_1}(dz_1,dz_2) \ .$$

Then

$$\begin{aligned}
|\phi_{X(t)}(u) - \phi_{X_\epsilon(t)}(u)| &= \phi_{X(t)}(u)|1 - \exp\{tf(u,\epsilon)\}| \\
&\leq \phi_{X(t)}(u)t|f(u,\epsilon)|\exp\{t|f(u,1)|\} \ .
\end{aligned}$$

With Taylor expansions we find that

$$\begin{aligned}
e^{iu(z_2-z_1)} &= 1 + iu(z_2-z_1) - \frac{1}{2}u^2(z_2-z_1)^2 \\
&\qquad + (iu(z_2-z_1))^3 \sum_{n=0}^{\infty} \frac{(iu(z_2-z_1))^n}{(n+3)!} \ ,
\end{aligned}$$

such that

$$\begin{aligned}
|f(u,\epsilon)| &= \left| \int_{|z|<\epsilon} (iu(z_2-z_1))^3 \sum_{n=0}^{\infty} \frac{(iu(z_2-z_1))^n}{(n+3)!} \nu_{\theta+\mathbf{1}_1}(dz_1,dz_2) \right| \\
&\leq \int_{|z|<\epsilon} |z_2-z_1|^3 |u|^3 \sum_{n=0}^{\infty} \frac{|u(z_2-z_1)|^n}{n!} \nu_{\theta+\mathbf{1}_1}(dz_1,dz_2) \\
&\leq \int_{|z|<\epsilon} |z_2-z_1|^3 |u|^3 \exp\{|u(z_2-z_1)|\} \nu_{\theta+\mathbf{1}_1}(dz_1,dz_2)
\end{aligned}$$

$$\leq c(u) \int_{|z|<\epsilon} |z_2 - z_1|^3 \nu_{\theta+\mathbf{1}_1}(dz_1, dz_2)$$

for $\epsilon < 1$ with

$$c(u) = |u|^3 \exp\{|u|\} ,$$

and the Proposition follows. □

As the integration is over $|z| < \epsilon$ we have

$$\begin{aligned}
&\int_{|z|<\epsilon} |z_2 - z_1|^3 \nu_{\theta+\mathbf{1}_1}(dz_1, dz_2) \\
&\quad = \int_{|z|<\epsilon} |z_2 - z_1||z_2 - z_1|^2 \nu_{\theta+\mathbf{1}_1}(dz_1, dz_2) \\
&\quad \leq 2\epsilon \int_{|z|<\epsilon} |z_2 - z_1|^2 \nu_{\theta+\mathbf{1}_1}(dz_1, dz_2) \\
&\quad = 2\epsilon \times \sigma_X^2(\epsilon) .
\end{aligned}$$

In terms of the variance of the original processes $L^{(1)}$ and $L^{(2)}$ (under the measure $\mathbb{Q}_{\theta+\mathbf{1}_1}$) it is

$$\begin{aligned}
\sigma_X^2(\epsilon) &= \int_{|z|<\epsilon} (z_2 - z_1)^2 \nu_{\theta+\mathbf{1}_1}(dz_1, dz_2) \\
&= \int_{|z|<\epsilon} z_2^2 \nu_{\theta+\mathbf{1}_1}(dz_1, dz_2) - 2\int_{|z|<\epsilon} z_2 z_1 \nu_{\theta+\mathbf{1}_1}(dz_1, dz_2) + \int_{|z|<\epsilon} z_1^2 \nu_{\theta+\mathbf{1}_1}(dz_1, dz_2) .
\end{aligned}$$

Using the Fourier representation of the option prices in the one-dimensional case, we get the following result

Proposition 3.3.3. *Assume that $\int_{\mathbb{R}} |\widehat{f}(iR - u)c(u - iR)|du < \infty$. Then it holds that*

$$|C_\epsilon - C| \leq d\epsilon\sigma_X^2(\epsilon) ,$$

for a constant d.

Again, if one chooses an approximation of X that only truncates the small jumps, but does not approximate them with a Brownian term, then analogously to the proof above we find an error estimate proportional to

$$d\sigma_X^2(\epsilon) .$$

3.3.2 Approximation under $\mathbb{P}$

Now, approximate under the measure $\mathbb{P}$. Like this we mimic the situation of two investors, who believe in different behaviours of the small variations. The first investor models them continuously by a Brownian motion, and the other by a Lévy process with infinite activity. That is, consider the bivariate price processes

$$S(t) = S(0)\mathrm{e}^{L(t)} ,$$

where the Lévy process $L = (L^{(1)}, L^{(2)})$ is given by (3.2.1) and

$$S_\epsilon(t) = S(0)\mathrm{e}^{L_\epsilon(t)} ,$$

where $L_\epsilon = (L_\epsilon^{(1)}, L_\epsilon^{(2)})$ is given by (3.2.2). Proposition 3.3.1 gives the prices for the spread options written on S and S_ϵ by

$$\begin{aligned} C &= S^{(1)}(0)\mathbb{E}_{\theta+\mathbf{1}_1}[f(X(T))] \\ C_\epsilon &= S^{(1)}(0)\mathbb{E}_{\theta_\epsilon+\mathbf{1}_1}[f(X_\epsilon(T))] , \end{aligned}$$

where $f = \max\Big((\mathrm{e}^x - 1), 0\Big)$ and

$$\begin{aligned} X(T) &= \widehat{a}T \\ &\quad + \int_0^T \int_{|z|<1} (z_2 - z_1)\widetilde{N}_{\theta+\mathbf{1}_1}(ds, dz) + \int_0^T \int_{|z|\geq 1} (z_2 - z_1)N_{\theta+\mathbf{1}_1}(ds, dz) \\ X_\epsilon(T) &= \widehat{a}_\epsilon T + (\alpha_2(\epsilon) - \alpha_1(\epsilon))B^{(1)}(T) + (\alpha_3(\epsilon) - \alpha_2(\epsilon))B^{(2)}(T) \\ &\quad + \int_0^T \int_{\epsilon\leq|z|<1} (z_2 - z_1)\widetilde{N}_{\theta_\epsilon+\mathbf{1}_1}(ds, dz) + \int_0^T \int_{|z|\geq 1} (z_2 - z_1)N_{\theta_\epsilon+\mathbf{1}_1}(ds, dz). \end{aligned}$$

Here, it is

$$\begin{aligned} \widehat{a} &= -\int_{\mathbb{R}_0^2} e^{z_2-z_1} - 1 - (z_2 - z_1)1_{|z|<1}\nu_{\theta+\mathbf{1}_1}(dz_1, dz_2) \\ \widehat{a}_\epsilon &= -\frac{1}{2}(\alpha_2(\epsilon) - \alpha_1(\epsilon))^2 - \frac{1}{2}(\alpha_3(\epsilon) - \alpha_2(\epsilon))^2 \\ &\quad - t\int_{\mathbb{R}_0^2} (e^{z_2-z_1} - 1 - (z_2 - z_1)1_{|z|<1})1_{|z|<\epsilon}\nu_{\theta_\epsilon+\mathbf{1}_1}(dz_1, dz_2) \end{aligned}$$

and $\alpha_i(\epsilon)$, $1 \leq i \leq 3$ are defined by equation (3.2.5). Then

$$\frac{1}{2}(\alpha_2(\epsilon) - \alpha_1(\epsilon))^2 + \frac{1}{2}(\alpha_3(\epsilon) - \alpha_2(\epsilon))^2 = \frac{1}{2}\int_{|z|<\epsilon} (z_2 - z_1)^2\nu(dz) .$$

Note that the Esscher transformation parameter θ_ϵ is defined such that the discounted price processes $\mathrm{e}^{-rt}S_\epsilon^{(i)}(t)$, $i = 1, 2$ are martingales, exactly in the same way as θ is defined such that $\mathrm{e}^{-rt}S^{(i)}(t)$, $i = 1, 2$ are martingales. Therefore, θ_ϵ it depends on ϵ. In the paper by Benth, Di Nunno, and Khedher [10] it is proved that θ_ϵ is bounded uniformly in ϵ in the case of a one-dimensional Lévy process. In our case it follows with the same arguments that θ_ϵ is bounded uniformly in ϵ and that

$$|\theta_i^\epsilon - \theta_i| \leq c_\theta\sigma_i^2(\epsilon), \qquad i = 1, 2, \tag{3.3.10}$$

where c_θ is a constant depending on θ and $\sigma_i^2(\epsilon)$, $i = 1, 2$, is given by (3.2.3).
We choose the scaling coefficients in (3.2.2) such that we keep the variance of the small jumps under $\mathbb{P}$. After a change of measure the Lévy measure changes and therefore also the

variation of the small jumps under the new measure. We find an estimate of the difference of the variation under the two different measures as follows.

$$\begin{aligned}
&\left|\int_{|z|<\epsilon}(z_2-z_1)^2\nu(dz)-\int_{|z|<\epsilon}(z_2-z_1)^2\nu_\theta(dz)\right| \qquad (3.3.11)\\
&\quad=\left|\int_{|z|<\epsilon}(z_2-z_1)^2(1-e^{\theta z})\nu(dz)\right|\\
&\quad\le\int_{|z|<\epsilon}(z_2-z_1)^2|\theta||z|e^{|\theta z|}\nu(dz)\\
&\quad\le c\epsilon\int_{|z|<\epsilon}(z_2-z_1)^2\nu(dz)
\end{aligned}$$

for a constant c.

Proposition 3.3.4. *It is*

$$\begin{aligned}
|\psi_{X(t)}(u)-\psi_{X_\epsilon(t)}(u)]| \le c_1(u)\epsilon\int_{|z|<\epsilon}(z_2-z_1)^2\nu(dz)+c_2(u)\int_{|z|<\epsilon}|z_2-z_1|^3\nu_{\theta+\mathbf{1}_1}(dz)\\
+c_3(u)\Big(\int_{|z|<\epsilon}z_1^2\nu(dz)+\int_{|z|<\epsilon}z_2^2\nu(dz)\Big)\,.
\end{aligned}$$

Proof. The characteristic functions of $X(t)$ and $X_\epsilon(t)$ under $\mathbb{Q}_{\theta_\epsilon+\mathbf{1}_1}$ and $\mathbb{Q}_{\theta+\mathbf{1}_1}$ resp. are given by

$$\begin{aligned}
\psi_{X(t)}(u)&=\exp\Big\{iu\widehat{a}t+t\int_{\mathbb{R}_0^2}\Big(\mathrm{e}^{iu(z_2-z_1)}-1-iu(z_2-z_1)1_{|z|<1}\Big)\nu_{\theta+\mathbf{1}_1}(dz)\Big\}\\
\psi_{X_\epsilon(t)}(u)&=\exp\Big\{iu\widehat{a}^\epsilon t-\frac{1}{2}u^2b_\epsilon t\\
&\qquad+t\int_{|z|\ge\epsilon}\Big(\mathrm{e}^{iu(z_2-z_1)}-1-iu(z_2-z_1)1_{|z|<1}\Big)\nu_{\theta_\epsilon+\mathbf{1}_1}(dz)\Big\}\,,
\end{aligned}$$

where $b_\epsilon=\int_{|z|<\epsilon}(z_2-z_1)^2\nu(dz)$. Let $f(u,\epsilon)$ be defined as follows

$$f(u,\epsilon)=\log\psi_{X_\epsilon(1)}(u)-\log\psi_{X(1)}(u)$$

$$\begin{aligned}
|f(u,\epsilon)|\le{}&|iu(\widehat{a}^\epsilon-\widehat{a})-\frac{1}{2}u^2b_\epsilon\\
&+\int_{|z|\ge\epsilon}\Big(\mathrm{e}^{iu(z_2-z_1)}-1-iu(z_2-z_1)1_{|z|<1}\Big)\nu_{\theta_\epsilon+\mathbf{1}_1}(dz)\\
&-\int_{\mathbb{R}_0^2}\Big(\mathrm{e}^{iu(z_2-z_1)}-1-iu(z_2-z_1)1_{|z|<1}\Big)\nu_{\theta+\mathbf{1}_1}(dz)\Big|\\
\le{}&+|iu(\widehat{a}^\epsilon-\widehat{a})|\\
&\left|-\frac{1}{2}u^2b_\epsilon-\int_{|z|<\epsilon}\big(\mathrm{e}^{iu(z_2-z_1)}-1-iu(z_2-z_1)\big)\mathrm{e}^{(\theta+\mathbf{1}_1).z}\nu(dz)\right|\\
&+\left|\int_{|z|\ge\epsilon}\big(\mathrm{e}^{iu(z_2-z_1)}-1-iu(z_2-z_1)1_{|z|<1}\big)\big(\mathrm{e}^{(\theta_\epsilon+\mathbf{1}_1).z}-\mathrm{e}^{(\theta+\mathbf{1}_1).z}\big)\nu(dz)\right|\,.
\end{aligned}$$

We have

$$\left|\mathrm{e}^{(\theta_\epsilon+\mathbf{1}_1).z} - \mathrm{e}^{(\theta+\mathbf{1}_1).z}\right| \leq (|\theta_1^\epsilon - \theta_1||z_1| + |\theta_2^\epsilon - \theta_2||z_2|)e^{|c||z|} \tag{3.3.12}$$

for a constant c. Then

$$\begin{aligned}
&\left|\int_{|z|\geq\epsilon}\left(\mathrm{e}^{iu(z_2-z_1)} - 1 - iu(z_2-z_1)1_{|z|<1}\right)\left(\mathrm{e}^{(\theta_\epsilon+\mathbf{1}_1).z} - \mathrm{e}^{(\theta+\mathbf{1}_1).z}\right)\nu(dz)\right| \\
&\quad\leq |\theta_1^\epsilon - \theta_1| \int_{|z|\geq\epsilon}\left|\mathrm{e}^{iu(z_2-z_1)} - 1 - iu(z_2-z_1)1_{|z|<1}\right| |z_1| e^{|c||z|}\nu(dz) \\
&\qquad + |\theta_2^\epsilon - \theta_2| \int_{|z|\geq\epsilon}\left|\mathrm{e}^{iu(z_2-z_1)} - 1 - iu(z_2-z_1)1_{|z|<1}\right| |z_2| e^{|c||z|}\nu(dz) \ .
\end{aligned}$$

As in the proof of Proposition 3.3.2 and using (3.3.11), we find

$$\begin{aligned}
&\left|-\frac{1}{2}u^2 b_\epsilon - \int_{|z|<\epsilon}\left(\mathrm{e}^{iu(z_2-z_1)} - 1 - iu(z_2-z_1)\right)\mathrm{e}^{(\theta+\mathbf{1}_1).z}\nu(dz)\right| \\
&\quad\leq c_1(u)\epsilon \int_{|z|<\epsilon}(z_2-z_1)^2\nu(dz) + c_2(u)\int_{|z|<\epsilon}|z_2-z_1|^3\nu_{\theta+\mathbf{1}_1}(dz) \ .
\end{aligned}$$

Thus we have

$$\begin{aligned}
|\psi_{X(t)}(u) - \psi_{X_\epsilon(t)}(u)]| &\leq |\psi_{X(t)}(u)||1 - \exp(f(u,\epsilon))| \\
&\leq \exp(tf(u,1))|\psi_{X(t)}(u)||tf(u,\epsilon)| \ .
\end{aligned}$$

Then the Proposition follows with (3.3.10). □

Then it follows as before

Proposition 3.3.5. *It holds that*

$$\begin{aligned}
|C - C_\epsilon| \leq d_1 \int_{|z|<\epsilon}|z_2-z_1|^3\nu_{\theta+\mathbf{1}_1}(dz) + d_2\epsilon\int_{|z|<\epsilon}(z_2-z_1)^2\nu(dz) + \\
+ d_3\Big(\int_{|z|<\epsilon} z_1^2\nu(dz) + \int_{|z|<\epsilon} z_2^2\nu(dz)\Big) \ .
\end{aligned}$$

The first term appears also in Proposition 3.3.3 and the second term results from the fact that we approximated before the measure change and so the approximation is not ”perfect” any more. Finally, the third term is due to the model dependency of the measure change and is an error estimation between the Esscher parameters θ and θ_ϵ.

CHAPTER 4

ARTICLE 3: "PRICING AND HEDGING OPTIONS IN ENERGY MARKETS BY BLACK-76"

Fred Espen Benth and Maren Diane Schmeck

Abstract

We prove that the price of options on forwards in commodity markets converge uniformly to the Black-76 formula when the short-term variations of the logarithmic spot price is a stationary Ornstein-Uhlenbeck process and the long-term variations are following a drifted Brownian motion. The convergence rate is exponential in the speed of mean-reversion and time to delivery of the underlying forward from the exercise time of the option. This can be applied to energy markets like electricity and gas to argue for the use of Black-76 in pricing of options, although the spot prices may show large spikes. Furthermore, we prove that the quadratic hedging strategy converges in a similar fashion to the delta-hedge in the Black-76 model. Our results are illustrated with a numerical example of relevance to energy markets.

4.1 Introduction

The typical stochastic models for spot and forward prices in oil, gas and electricity markets separate the time evolution into long-term and short-term factors. The long-term effects include inflation and depletion of reserves (in case of non-renewable commodities like oil and gas), and is typically thought of as being non-stationary. On the other hand, the prices are shocked by short-term effects like outages of power plants or changes in demand from temperature variations. These effects are modelled by stationary, mean-reverting processes. The classical model of Gibson and Schwartz [42] defines the logarithmic oil spot prices as a drifted Brownian motion and an Ornstein-Uhlenbeck process. This two-factor model, consisting of a non-stationary and a stationary part, has later been applied to gas and electricity markets (see e.g. Lucia and Schwartz [47]), in particular to forward pricing.

In energy markets like gas and power, options traded on exchanges are typically written on futures contracts delivering the underlying energy over a specified period. For example, at the Nordic power market NordPool call and put options are traded based on futures contract with financial delivery of electricity over given months. One may suspect that the short-term factor of the forward price evolution inherited from the spot will be insignificant in the option price. Due to the delivery period, short-term shocks in the spot may vanish in the futures dynamics due to smoothing by the delivery period. This means that one is left with the non-stationary part, which leads to the claim that the option price can be approximated well by the Black-76 formula. In this paper we show that this is indeed true in many practically relevant situations. In fact, we prove a uniform exponential convergence of the "true" option price towards the one given by Black-76 in terms of the speed of mean reversion of the short-term stationary factor and the time left to delivery of the underlying futures from the exercise time of the option. A "folklore" in the NordPool market says that one can do well with Black-76. We show that this is indeed the case, justified by theoretical results and numerical examples.

In Lucia and Schwartz [47], the forward price dynamics for contracts delivering electricity over a specified period is defined as the average of forwards with fixed delivery time. As the spot model is defined as an exponential process, there exists no analytic formula for the forward price delivering over a period for models of interest. In this paper we view this differently, and think of the forward price with delivery period as a contract with "fixed-delivery" given by the mid-point of the delivery period. In this way we can make a reasonable approximation of the forward price dynamics in electricity and gas, where one does not have the classical convergence of forward price to spot when time to start of delivery goes to zero (see Benth et al. [19] for more details). Obviously, for other commodities (like oil), where the forward delivers at a given time we do not need to use such an approach.

A typical characteristic of gas and electricity markets are sudden large price deviations, frequently referred to as spikes. For example, the German power market EEX shows a significant amount of negative price spikes, mainly due to wind power generation. More usual are the positive spikes, which for example can be seen in the NordPool market during winter season. Also in gas markets one sees large price fluctuations occurring due to for example cold weather (see for example Geman [40] for a discussion). These big price fluctuations call for models based on non-Gaussian stochastic drivers, and the application of Lévy processes, possibly time-inhomogeneous, seems natural (see Benth et al. [19]). In this paper we model the short-term dynamics by an Ornstein-Uhlenbeck process driven by a Lévy process. In this way we can include modelling of spikes, or large variations, in the price dynamics.

The implication of a spot price driven by Lévy innovations is that the forward price dynamics become more involved. Also, we are put in an incomplete market setting which makes pricing and hedging of the option a delicate problem. As we choose to introduce a pricing measure Q based on the Girsanov and Esscher transforms (see Benth et al. [19]), we have already pinned down a risk-neutral probability for the forward prices (namely the one we choose when deriving the forward from the spot). We can then derive the call option price based on a conditional expectation of the payout from the call. On the other hand, there exists no hedging strategy perfectly replicating the option.

There exists many approaches to hedging in incomplete markets, where one soughts to find a strategy in the underlying which minimizes the risk exposed in a short position of the option (see Cont and Tankov [31]). We focus here on the quadratic hedging strategy, which minimizes the L^2-distance between the payout from the option and the hedging portfolio.

We refer to Cont and Tankov [31] for more on this strategy in incomplete markets where the underlying asset price is defined as an exponential Lévy process. We are able to determine the quadratic hedging strategy for our market model, and express this in terms of the option price and its sensitivity to the underlying (the delta). As it turns out, we are able to show that the quadratic hedge converges uniformly to the simple delta-hedging strategy, and moreover, we determine the rate of convergence to be the same as for the price, namely exponential in time to delivery and mean-reversion speed.

Our findings are presented as follows. In the next Section we introduce our spot price model and derive the forward price dynamics. Section 3 deals with the convergence of option prices towards the Black-76. This Section also presents a Fourier-based pricing formula as well as a numerical illustration. Next, in Section 4, we derive the quadratic hedging strategy for call options on forwards, and prove that this converges exponentially to the delta hedge of Black-76.

4.2 The spot and forward price dynamics

Fix a filtered complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$, and suppose that the energy spot price follows a two factor model defined as

$$S(t) = \Lambda(t)\exp(X(t) + Y(t)) \,. \tag{4.2.1}$$

Here, the non-stationary factor X is a drifted Brownian motion

$$dX(t) = \mu\, dt + \sigma\, dB(t) \,, \tag{4.2.2}$$

with B being a Brownian motion and $\mu, \sigma > 0$ constants. The stationary factor Y is given by the Ornstein-Uhlenbeck dynamics

$$dY(t) = -\beta Y(t)\, dt + dL(t) \,, \tag{4.2.3}$$

where L is a pure jump Lévy process with Lévy-Khintchine decomposition

$$L(t) = \int_0^t \int_{|z|<1} z\tilde{N}(ds, dz) + \int_0^t \int_{|z|\geq 1} zN(dt, dz) \,,$$

and $\beta > 0$ a constant. The deterministic seasonality function $\Lambda(t) : \mathbb{R}_+ \to \mathbb{R}_+$ is supposed to be continuous.

The exponential two-factor dynamics (4.2.1) is a generalization of the spot price model proposed by Gibson and Schwartz [42] (see also Schwartz and Smith [55]). They assumed L to be a Brownian motion correlated with B, and applied the model to a study of oil spot and forward prices. Later, Lucia and Schwartz [47] suggested such a two factor model for electricity spot and forward prices, again using L as a Brownian motion. They studied empirically NordPool data. The two-factor model takes into account mean reversion of the commodity price as well as uncertainty in the equilibrium level to which the prices revert. The non-stationary long time factor models the equilibrium price level, and reflects expectations on for example improving technologies for the production of the commodity, inflation or political and regulatory effects, and depletion of non-renewable resources like gas and coal. The mean reverting short term factor describes changes in demand and supply resulting for example from variations in the weather conditions and sudden outages of power plants. They are tempered by the ability of market participants to respond to the changing market conditions and are therefore reverting back to their mean level. Lucia and Schwartz [47] provide evidence that one finds seasonal regular patterns in the electricity spot prices, accounted for in the model by the function Λ.

To make our analysis slightly simpler, we shall assume that L and B are independent. Moreover, as already stated, we let L be a pure-jump Lévy process and denote its Lévy measure by $\ell(dz)$. The motivation behind assuming a Lévy process rather than a Brownian motion driving the stationary part comes from power markets, where the spot prices are known to have spikes. Such spikes are typically of short duration, and can be reasonably well modelled by a jump (in the Lévy process) followed by fast mean reversion (coming from a large β). Also in gas markets one expects big short term variations, where a Lévy process seems more natural to span the uncertainty than a Brownian motion driven Ornstein-Uhlenbeck process. We refer to Benth et al. [19] for more discussions motivating the use of jump processes in energy markets.

We assume that the Lévy process has finite exponential moments, that is,

$$\int_1^\infty e^{cz}\,\ell(dz) + \int_{-\infty}^{-1} e^{-cz}\,\ell(dz) < \infty\,, \tag{4.2.4}$$

for a positive constant c. As we shall see, we need to have the constant $c \geq 3$ in order to prove our results. Hence, we suppose that this is true from now on. Finally, we denote by ϕ the logarithmic moment generating function of $L(1)$, defined as

$$\phi(\theta) = \ln \mathbb{E}[\exp(\theta L(1))]\,,$$

which exists for $|\theta| \leq 3$.

Since our attention is on pricing call option written on forward contracts, we need to relate the forward price dynamics to the spot model. The standard definition of the forward price $f(t,T)$ at time $t \geq 0$ of a contract delivering the underlying energy at time $T \geq t$ is

$$f(t,T) = \mathbb{E}_Q[S(T)\,|\,\mathcal{F}_t] \tag{4.2.5}$$

for some pricing measure Q being equivalent to P. We implicitly assume here that $S(T)$ is integrable with respect to the pricing measure Q. In electricity, say, the spot is not storable, and *any* equivalent measure Q can be used as a pricing measure (see Benth et al. [19]). Gas can be stored and traded in a spot market, but transportation and storage costs will be incurred. The same is the case of oil. In addition, one talks about the convenience yield for these commodities. Collected together, one may view the storage costs, transportation and convenience yield as a result of a measure change, or, vice versa, that a measure change from P to Q is a modelling of these three components. Thus, also in the gas and oil situation, it is convenient to define a rich class of equivalent probability measures which can flexibly model the drift imposed by storage, transportation and convenience yield. The standard class of probabilities is provided by the Esscher transform, which coincides with the Girsanov transform for the Brownian motion case. Using a constant Esscher transform (see Benth et al. [19]), the effect on the stationary factor is an additional drift coefficient adding on the μ, and for the Lévy process the effect will be an exponential tilting of the Lévy measure, but preserving the Lévy property. Hence, in order to keep notation at a minimum, we suppose that our spot model is already stated under a pricing measure Q (or, we can just re-interpret the meaning of the coefficients in the spot model).

The next proposition states the forward price explicitly in terms of the logarithmic moment generating function of $L(1)$.

Proposition 4.2.1. *The forward price $f(t,T)$ at $t \geq 0$ with delivery at $T \geq t$ is*

$$f(t,T) = h(t,T)\exp\left(X(t) + \mathrm{e}^{-\beta(T-t)}Y(t)\right)$$

with

$$h(t,T) = \Lambda(T)\exp\left(\mu(T-t) + \frac{1}{2}\sigma^2(T-t) + \int_t^T \phi(\mathrm{e}^{-\beta(T-s)})\,ds\right).$$

Proof. First, notice that

$$X(T) = X(t) + \mu(T-t) + \sigma(B(T) - B(t)),$$

and

$$Y(T) = \mathrm{e}^{-\beta(T-t)}Y(t) + \int_t^T \mathrm{e}^{-\beta(T-s)}\,dL(s)$$

by a straightforward use of the Itô formula for jump processes. But then, by the $\mathcal{F}_t$-adaptedness of $X(t)$ and $Y(t)$, the independent increment property of Lévy processes and the independence between B and L, we find

$$\begin{aligned} f(t,T) &= \Lambda(T)\mathbb{E}[\exp(X(T) + Y(T)) \,|\, \mathcal{F}_t] \\ &= \Lambda(T)\exp\left(\mu(T-t) + X(t) + \mathrm{e}^{-\beta(T-t)}Y(t)\right)\mathbb{E}\left[\exp(\sigma(B(T) - B(t))\right] \\ &\quad \times \mathbb{E}\left[\exp(\int_t^T \mathrm{e}^{-\beta(T-s)}\,dL(s))\right] \\ &= h(t,T)\exp\left(X(t) + \mathrm{e}^{-\beta(T-t)}Y(t)\right). \end{aligned}$$

This proves the result. □

We can find the dynamics of the forward price, which shows that it is indeed a geometric jump-diffusion model:

Proposition 4.2.2. *The dynamics of the process $t \mapsto f(t,T)$ for $t \leq T$ is*

$$\frac{df(t,T)}{f(t-,T)} = \sigma\,dB(t) + \int_{\mathbb{R}} \left\{\exp\left(z\mathrm{e}^{-\beta(T-t)}\right) - 1\right\}\widetilde{N}(dz,dt),$$

where $\widetilde{N}(dt,dz)$ is the compensated Poisson random measure of L and $f(t-,T)$ denotes the left-limit of $f(t,T)$.

Proof. Observe that f has finite expectation using (4.2.4) and that by definition, $t \mapsto f(t,T)$ is a martingale. This information simplifies considerably the application of Itô's Formula for jump processes, which shows the result. □

We remark that there are several papers modelling forward prices in energy directly rather than as a derivative of the spot price dynamics. A direct modelling of forward prices, following the so-called Heath-Jarrow-Morton approach from interest rate theory, has been extensively discussed in Benth et al. [19], as well as Benth and Koekebakker [17]. One natural class of such models may in fact be the dynamical model stated in Prop. 4.2.2.

4.3 Pricing call options on forwards

With the forward price at hand we go on and analyse the price of options on forwards. We focus our attention on European call options, and remark that put options can be priced via the call-put parity (see Benth et al. [19]).

To this end, we let $\tau \leq T$ be the exercise time of the call option, with a strike price $K > 0$. To simplify the exposition slightly, we assume that the risk-free interest rate is equal to zero, that is, $r = 0$. The no-arbitrage price of a call option at time $t \leq \tau$ written on a forward contract with price dynamics given as in Prop. 4.2.1, or equivalently Prop. 4.2.2, is defined by

$$C(t, \tau, T) = \mathbb{E}\left[\max(f(\tau, T) - K, 0) \,|\, \mathcal{F}_t\right] .$$

By Prop. 4.2.2, we see that the forward price is Markovian, and hence we find that the price of the call can be expressed as $C(t, \tau, T, f(t, T))$, with $C(t, \tau, T, x)$ given by

$$C(t, \tau, T, x) = \mathbb{E}\left[\max(f(\tau, T) - K, 0) \,|\, f(t, T) = x\right] .$$

Our aim now is to analyse this price in relation to the Black-76 formula. For the convenience of the reader, we have stated this famous formula for the price of a call option written on a forward with a geometric Brownian motion dynamics (see Black [21]).

Proposition 4.3.1. *Suppose the forward price dynamics is a geometric Brownian motion*

$$\frac{df(t, T)}{f(t, T)} = \sigma\, dB(t) .$$

Then the price at time t of a call option with strike K and exercise time $t \leq \tau \leq T$, is given by $C_{B76}(t, f(t, T))$ with

$$C_{B76}(t, \tau, T, x) = x\Phi(d_1(x)) - K\Phi(d_2(x))$$

for Φ being the cumulative standard normal distribution function, and

$$\begin{aligned} d_1(x) &= d_2 + \sigma\sqrt{\tau - t} \\ d_2(x) &= \frac{\ln\left(\frac{x}{K}\right) - \frac{1}{2}\sigma^2(\tau - t)}{\sigma\sqrt{\tau - t}} . \end{aligned}$$

Proof. See Black [21]. □

We want to show that $C(t, \tau, T, x)$ is converging to $C_{\text{B76}}(t, \tau, T, x)$ as the delivery time T of the underlying forward goes to infinity. Moreover, we want to have the rate of convergence measured in terms of the speed of mean reversion β of the spike component.

The price $C(t, \tau, T, x)$ can be represented as follows:

Proposition 4.3.2. *The price of a call option on the forward given in Prop. 4.2.1 is*

$$C(t,\tau,T,x) = x\mathbb{E}\Big[\exp\left(\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s) - \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds\right) \\ \times \Phi\left(d_1\left(x, \int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)\right)\right)\Big] \\ - K\mathbb{E}\left[\Phi\left(d_2\left(x, \int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)\right)\right)\right]$$

where $\phi(x)$ is the logarithmic moment generating function of $L(1)$ and

$$d_1(x,v) = d_2(x,v) + \sigma\sqrt{\tau - t}$$
$$d_2(x,v) = \frac{\ln\left(\frac{x}{K}\right) + v - \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds - \frac{1}{2}\sigma^2(\tau-t)}{\sigma\sqrt{\tau-t}}\,.$$

Proof. First, from Prop. 4.2.1, we have

$$f(\tau,T) = h(\tau,T)\exp\left(X(\tau) + \mathrm{e}^{-\beta(T-\tau)}Y(\tau)\right)$$
$$= f(t,T)\frac{h(\tau,T)}{h(t,T)}\exp\left(X(\tau) - X(t) + \mathrm{e}^{-\beta(T-\tau)}Y(\tau) - \mathrm{e}^{-\beta(T-t)}Y(t)\right)\,.$$

But,

$$\mathrm{e}^{-\beta(T-\tau)}Y(\tau) = Y(t)\mathrm{e}^{-\beta(T-t)} + \int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)$$

and

$$X(\tau) - X(t) = \mu(\tau - t) + \sigma(B(\tau) - B(t))\,.$$

Furthermore,

$$\frac{h(\tau,T)}{h(t,T)}\mathrm{e}^{\mu(\tau-t)} = \exp\left(-\frac{1}{2}\sigma^2(\tau-t) - \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds\right)\,.$$

Hence,

$$f(\tau,T) = f(t,T)\exp\left(\sigma(B(\tau)-B(t)) - \frac{1}{2}\sigma^2(\tau-t)\right) \\ \times \exp\left(\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s) - \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds\right)\,.$$

Denote by $Z(x)$ the random variable

$$Z(x) = x\exp\left(\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s) - \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds\right)\,,$$

and since L is independent of B, $Z(x)$ is independent of $B(\tau) - B(t)$. Conditioning on Z, yields

$$C(t,\tau,T,x)$$

$$
\begin{aligned}
&= \mathbb{E}\left[\max(f(\tau,T)-K,0)\,|\,x=f(t,T)\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\max\left(Z(x)\exp\left(\sigma(B(\tau)-B(t))-\frac{1}{2}\sigma^2(\tau-t)\right)-K,0\right)\,|\,Z(x)\right]\right].
\end{aligned}
$$

The inner expectation can be computed by the Black-76 formula in Prop. 4.3.1, with $Z(x)$ playing the role of x. Hence, the result follows. □

The expression for the price in the proposition above can now be used to show the convergence to the Black-76 formula. We prove this by a sequence of Lemmas. But first, let us introduce the log-moment generating function

$$
\phi_\beta(\theta) := \ln \mathbb{E}\left[\exp\left(\int_t^\tau \theta \mathrm{e}^{-\beta(T-s)}\,dL(s)\right)\right].
$$

From Benth et al. [19] we have

$$
\phi_\beta(\theta) = \int_t^\tau \phi\left(\theta \mathrm{e}^{-\beta(T-s)}\right)\,ds\,,
$$

for ϕ being the log-moment generating function of $L(1)$. Observe that $\phi_\beta(\theta)$ is well-defined for all $|\theta| \leq 3$. In the proof of the convergence to the Black-76 formula, we will need the following simple result.

Lemma 4.3.3. *The function $f(x) = (1-\exp(-x))/x$ for $x \geq 0$ is decreasing to zero with $f(0) = 1$.*

Proof. By L'Hopital's rule we find $f(0) = 1$. Moreover,

$$
f'(x) = \frac{(x+1)\mathrm{e}^{-x}-1}{x^2}\,,
$$

and since $x+1 \leq \mathrm{e}^x$ it holds that $(x+1)\mathrm{e}^{-x}-1 \leq 0$ and thus $f'(x) \leq 0$. Letting $x \to \infty$, we see that $f(x) \to 0$. The Lemma holds. □

In the results below, the positive constant c will be generic and not necessarily refer to the same value. We have,

Lemma 4.3.4. *It holds, for $\tau \leq T$,*

$$
\sup_{x\geq 0}\left|\mathbb{E}\left[\Phi\left(d_2\left(x,\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)\right)\right)\right]-\Phi(d_2(x,0))\right| \leq c\mathrm{e}^{-\beta(T-\tau)}
$$

for a constant $c > 0$.

Proof. It holds,

$$
\begin{aligned}
&\left|\mathbb{E}\left[\Phi\left(d_2\left(x,\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)\right)\right)\right]-\Phi(d_2(x,0))\right| \\
&\qquad \leq \mathbb{E}\left[\left|\Phi\left(d_2\left(x,\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)\right)\right)-\Phi(d_2(x,0))\right|\right].
\end{aligned}
$$

By the mean value theorem, there exists a random variable Z such that

$$\begin{aligned}\Phi & \left(d_2\left(x, \int_t^\tau \mathrm{e}^{-\beta(T-s)}\, dL(s)\right)\right)-\Phi(d_2(x,0)) \\ & =\Phi'(Z)\left(d_2\left(x, \int_t^\tau \mathrm{e}^{-\beta(T-s)}\, dL(s)\right)-d_2(x,0)\right).\end{aligned}$$

But, from the definition of Φ,

$$\Phi'(Z)=\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-Z^2/2}\leq\frac{1}{\sqrt{2\pi}}<1\,.$$

Furthermore,

$$d_2\left(x, \int_t^\tau \mathrm{e}^{-\beta(T-s)}\, dL(s)\right)=d_2(x,0)+\frac{1}{\sigma\sqrt{\tau-t}}\int_t^\tau \mathrm{e}^{-\beta(T-s)}\, dL(s)\,.$$

We therefore find, after using Cauchy-Schwarz' inequality,

$$\begin{aligned}&\left|\mathbb{E}\left[\Phi\left(d_2\left(x, \int_t^\tau \mathrm{e}^{-\beta(T-s)}\, dL(s)\right)\right)\right]-\Phi(d_2(x,0))\right| \\ &\qquad < \frac{1}{\sigma\sqrt{\tau-t}}\mathbb{E}\left[\left|\int_t^\tau \mathrm{e}^{-\beta(T-s)}\, dL(s)\right|\right] \\ &\qquad \leq \frac{1}{\sigma\sqrt{\tau-t}}\mathbb{E}\left[\left(\int_t^\tau \mathrm{e}^{-\beta(T-s)}\, dL(s)\right)^2\right]^{1/2}.\end{aligned}$$

From basic probability theory, we find

$$\begin{aligned}&\mathbb{E}\left[\left(\int_t^\tau \mathrm{e}^{-\beta(T-s)}\, dL(s)\right)^2\right] \\ &\qquad = \frac{d^2}{d\theta^2}\mathrm{e}^{\phi_\beta(\theta)}\Big|_{\theta=0} \\ &\qquad = \phi''_\beta(0)+(\phi'_\beta(0))^2 \\ &\qquad = \phi''(0)\int_t^\tau \mathrm{e}^{-2\beta(T-s)}\, ds+(\phi'(0))^2(\int_t^\tau \mathrm{e}^{-\beta(T-s)}\, ds)^2 \\ &\qquad = \left(\phi''(0)\frac{1}{2\beta}(1-\mathrm{e}^{-2\beta(\tau-t)})+(\phi'(0))^2\frac{1}{\beta^2}(1-e^{-\beta(\tau-t)})^2\right)\mathrm{e}^{-2\beta(T-\tau)}\,,\end{aligned}$$

But then we have

$$\begin{aligned}&\left|\mathbb{E}\left[\Phi\left(d_2\left(x, \int_t^\tau \mathrm{e}^{-\beta(T-s)}\, dL(s)\right)\right)\right]-\Phi(d_2(x,0))\right| \\ &\qquad \leq \frac{1}{\sigma}\left[\phi''(0)\frac{1-\mathrm{e}^{-2\beta(\tau-t)}}{2\beta(\tau-t)}+(\phi'(0))^2\frac{1}{\beta^2}\frac{(1-\mathrm{e}^{-\beta(\tau-t)})^2}{\tau-t}\right]^{\frac{1}{2}}\mathrm{e}^{-\beta(T-\tau)}\,.\end{aligned}$$

Since $1-\exp(-\beta(\tau-t))\leq 1$, we use Lemma 4.3.3 twice to conclude the proof. □

In our next Lemma, we estimate the difference between $\Phi(d_2(x,0))$ and $\Phi(d_2(x))$.

Lemma 4.3.5. *It holds, for $\tau \leq T$,*

$$\sup_{x\geq 0} |\Phi(d_2(x,0)) - \Phi(d_2(x))| \leq c\mathrm{e}^{-\beta(T-\tau)},$$

for a constant $c > 0$.

Proof. We have that

$$d_2(x,0) = d_2(x) - \frac{1}{\sigma\sqrt{\tau - t}} \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\, ds .$$

But then, appealing to the mean value theorem,

$$|\Phi(d_2(x,0)) - \Phi(d_2(x))| \leq \frac{c}{\sigma\sqrt{\tau - t}} \left| \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\, ds \right| .$$

We analyse the integral on the right-hand side in more detail. For notational simplicity, let $\gamma(s) = \exp(-\beta(T-s))$. By definition of the log-moment generating function

$$\int_t^\tau \phi(\gamma(s))\, ds = \int_t^\tau \int_{\mathbb{R}} \left\{ \mathrm{e}^{\gamma(s)z} - 1 - \gamma(s)z\mathbf{1}_{|z|<1} \right\} \ell(dz)\, ds .$$

We have for $|z| \geq 1$,

$$\begin{aligned} |\mathrm{e}^{\gamma(s)z} - 1| &\leq \sum_{n=1}^\infty \frac{(\gamma(s)|z|)^n}{n!} \\ &= \gamma(s)|z| \sum_{n=1}^\infty \frac{(\gamma(s)|z|)^{n-1}}{n!} \\ &\leq \gamma(s)|z| \sum_{n=0}^\infty \frac{(\gamma(s)|z|)^n}{n!} \\ &= \gamma(s)|z|\mathrm{e}^{\gamma(s)|z|} . \end{aligned}$$

If $|z| < 1$, the series representation of the exponential function gives

$$\begin{aligned} |\mathrm{e}^{\gamma(s)z} - 1 - \gamma(s)z| &\leq \sum_{n=2}^\infty \frac{(\gamma(s)|z|)^n}{n!} \\ &\leq \gamma^2(s)|z|^2 \sum_{n=2}^\infty \frac{\gamma^{n-2}(s)}{n!} \\ &= \gamma^2(s)|z|^2\mathrm{e}^{\gamma(s)} . \end{aligned}$$

Hence, using the definition of $\gamma(s)$,

$$\int_{\mathbb{R}} |\mathrm{e}^{\gamma(s)z} - 1 - \gamma(s)z\mathbf{1}_{|z|<1}|\, \ell(dz)$$

$$
\begin{aligned}
&\leq \gamma^2(s) \int_{|z|<1} z^2 \, \ell(dz) \mathrm{e}^{\gamma(s)} + \gamma(s) \int_{|z|\geq 1} |z| \mathrm{e}^{\gamma(s)|z|} \, \ell(dz) \\
&\leq \mathrm{e}^{-2\beta(T-s)} \int_{|z|<1} z^2 \, \ell(dz) + \mathrm{e}^{-\beta(T-s)} \int_{|z|\geq 1} \mathrm{e}^{2|z|} \, \ell(dz) \\
&\leq \mathrm{e}^{-\beta(T-s)} \left(\int_{|z|<1} z^2 \, \ell(dz) + \int_{|z|\geq 1} \mathrm{e}^{2|z|} \, \ell(dz) \right) .
\end{aligned}
$$

Therefore, there exists a constant $c > 0$ such that

$$
\begin{aligned}
\left| \frac{1}{\sigma\sqrt{\tau - t}} \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)}) \, ds \right| &\leq \frac{c}{\sigma\sqrt{\tau - t}} \int_t^\tau \mathrm{e}^{-\beta(T-s)} \, ds \\
&= \frac{c}{\sigma\beta} \frac{1 - \mathrm{e}^{-\beta(\tau - t)}}{\sqrt{\tau - t}} \mathrm{e}^{-\beta(T-\tau)} \, .
\end{aligned}
$$

The Lemma follows by invoking Lemma 4.3.3. □

We move on analysing the first term in our pricing formula in Prop. 4.3.2.

Lemma 4.3.6. *It holds, for $\tau \leq T$,*

$$
\begin{aligned}
\sup_{x\geq 0} \Big| \mathbb{E}\Big[\exp\Big(\int_t^\tau \mathrm{e}^{-\beta(T-s)} \, dL(s) - \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)}) \, ds \Big) \\
\times \Phi\Big(d_1\Big(x, \int_t^\tau \mathrm{e}^{-\beta(T-s)} \, dL(s) \Big) \Big) \Big] - \Phi(d_1(x, 0)) \Big| \leq c\mathrm{e}^{-\beta(T-\tau)}
\end{aligned}
$$

for a constant $c > 0$.

Proof. Introduce the probability $\widetilde{Q}$ with Radon-Nikodym derivative

$$
\frac{d\widetilde{Q}}{dQ} = \exp\left(\int_t^\tau \mathrm{e}^{-\beta(T-s)} \, dL(s) - \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)}) \, ds \right) .
$$

This is an Esscher transform, turning the Lévy process L into a independent increment process (with time-dependent compensator measure, see Benth et al. [19]). The logarithmic-moment generating function of $\int_t^\tau \mathrm{e}^{-\beta(T-s)} \, dL(s)$ under $\widetilde{Q}$ will become

$$
\begin{aligned}
\phi_{\beta,\widetilde{Q}}(\theta) &= \ln \mathbb{E}_{\widetilde{Q}} \left[\exp\left(\theta \int_t^\tau \mathrm{e}^{-\beta(T-s)} \, dL(s) \right) \right] \\
&= \ln \mathbb{E} \left[\exp\left(\int_t^\tau \{(1+\theta)\mathrm{e}^{-\beta(T-s)}\} \, dL(s) \right) \right] - \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)}) \, ds \\
&= \int_t^\tau \{ \phi((1+\theta)\mathrm{e}^{-\beta(T-s)}) - \phi(\mathrm{e}^{-\beta(T-s)}) \} \, ds \, .
\end{aligned}
$$

Since

$$
d_1(x, v) = d_1(x, 0) + \frac{v}{\sigma\sqrt{\tau - t}} \, ,
$$

we estimate as in the proof of Lemma 4.3.4 to get

$$\Big|\mathbb{E}\left[\exp\left(\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)-\int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds\right)\Phi\left(d_1\left(x,\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)\right)\right)\right]$$
$$-\Phi(d_1(x,0))\Big| \le \frac{1}{\sigma\sqrt{\tau-t}}\mathbb{E}_{\widetilde{Q}}\left[\left(\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)\right)^2\right]^{1/2}.$$

Again, from elementary probability theory we find

$$\mathbb{E}_{\widetilde{Q}}\left[\left(\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)\right)^2\right]=\phi''_{\beta,\widetilde{Q}}(0)+(\phi'_{\beta,\widetilde{Q}}(0))^2\,.$$

From our definition of $\phi_{\beta,\widetilde{Q}}$, we find by appealing to the dominated convergence theorem,

$$\phi'_{\beta,\widetilde{Q}}(\theta)=\int_t^\tau \phi'((1+\theta)\mathrm{e}^{-\beta(T-s)})\mathrm{e}^{-\beta(T-s)}\,ds\,,$$

which implies

$$\phi'_{\beta,\widetilde{Q}}(0)=\int_t^\tau \phi'(\mathrm{e}^{-\beta(T-s)})\mathrm{e}^{-\beta(T-s)}\,ds\,.$$

Denoting $\gamma(s)=\exp(-\beta(T-s))$, it follows,

$$\begin{aligned}\phi'(\gamma(s))&=\frac{d}{d\xi}\int_{\mathbb{R}}\{\mathrm{e}^{\xi z}-1-\xi z\mathbf{1}_{|z|<1}\}\,\ell(dz)\Big|_{\xi=\gamma(s)}\\&=\int_{\mathbb{R}}\{z\mathrm{e}^{\gamma(s)z}-z\mathbf{1}_{|z|<1}\}\,\ell(dz)\\&=\int_{|z|<1}z\{\mathrm{e}^{\gamma(s)z}-1\}\,\ell(dz)+\int_{|z|\ge 1}z\mathrm{e}^{\gamma(s)z}\,\ell(dz)\,.\end{aligned}$$

As

$$|\mathrm{e}^{\gamma(s)z}-1|\le |z|\mathrm{e}^{\gamma(s)|z|}\le |z|\mathrm{e}^1\,,$$

for $|z|<1$, while for $|z|>1$

$$|z|\mathrm{e}^{\gamma(s)|z|}\le \mathrm{e}^{2|z|}\,,$$

it follows that

$$|\phi'(\gamma(s))|\le \mathrm{e}^1\int_{|z|<1}z^2\,\ell(dz)+\int_{|z|\ge 1}\mathrm{e}^{2|z|}\,\ell(dz)\le c\,,$$

for a positive constant c. Similarly,

$$\begin{aligned}\phi''(\gamma(s))&=\frac{d}{d\xi}\int_{\mathbb{R}}z\{\mathrm{e}^{\xi z}-\mathbf{1}_{|z|<1}\}\,\ell(dz)\Big|_{\xi=\gamma(s)}\\&=\int_{\mathbb{R}}z^2\mathrm{e}^{\gamma(s)z}\,\ell(dz)\,.\end{aligned}$$

As

$$z^2 \mathrm{e}^{\gamma(s)|z|} \le \mathrm{e}^1 |z|^2 \mathbf{1}_{|z|<1} + \mathbf{1}_{|z|\ge 1} \mathrm{e}^{3|z|} \,,$$

it follows from the condition on the Lévy measure in (4.2.4)

$$|\phi''(\mathrm{e}^{-\beta(T-s)})| \le c$$

for some constant $c > 0$. Wrapping up the estimates, we are left with

$$\frac{1}{\sigma\sqrt{\tau - t}} \mathbb{E}_{\tilde{Q}} \left[\left(\int_t^\tau \mathrm{e}^{-\beta(T-s)} \, dL(s) \right)^2 \right]^{1/2}$$
$$\le \frac{1}{\sigma} \left[c \, \frac{1}{\beta^2} \frac{(1 - e^{-\beta(\tau - t)})^2}{\tau - t} + c \, \frac{1 - \mathrm{e}^{-2\beta(\tau - t)}}{\tau - t} \right]^{\frac{1}{2}} \mathrm{e}^{-\beta(T-\tau)}$$

As the fractions in the last inequality by Lemma 4.3.3 are bounded, the Lemma is proven. □

We end with the Lemma,

Lemma 4.3.7. *It holds, for $\tau \le T$,*

$$\sup_{x \ge 0} |\Phi(d_1(x, 0)) - \Phi(d_1(x))| \le c\mathrm{e}^{-\beta(T-\tau)} \,,$$

for some constant $c > 0$.

Proof. Since

$$d_1(x, 0) = d_1(x) - \frac{1}{\sigma\sqrt{\tau - t}} \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)}) \, ds \,,$$

the proof is identical to the proof of Lemma 4.3.5. □

We summarize our findings in the following theorem:

Theorem 4.3.8. *Suppose that $\tau \le T$. Then it holds that*

$$\sup_{x \ge 0} |C(t, \tau, T, x) - C_{B76}(t, \tau, T, x)| \le c\mathrm{e}^{-\beta(T-\tau)} \,,$$

for some constant c.

Proof. Appealing to the triangle inequality and Lemmas 4.3.4-4.3.7 yield the result. □

By fixing τ, we see that the call option price $C(t, \tau, T, x)$ is converging uniformly to the Black-76 price as $T \to \infty$. The convergence is exponential with the rate β. We recall that β is the speed of mean reversion of the spike factor of the spot dynamics. Note also that tracing through the proofs of the Lemmas 4.3.4-4.3.7 we can find an expression for the constant c, and therefore we can find the maximal error between the Black-76 price and the ”correct”

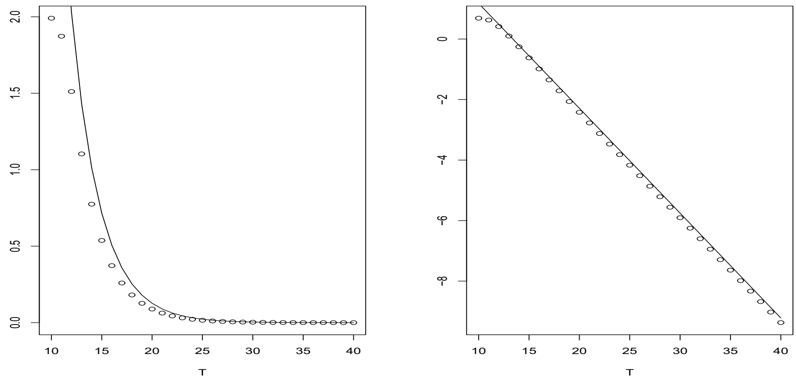

Figure 4.1: Difference of the option price to the Black-76 price (left), and on log-scale (right). The solid line is the theoretical error estimate.

price $C(t, \tau, T, x)$. However, the overall conclusion is that for options where the delivery time T is sufficiently bigger than τ, the call option price can be approximated with a high degree of accuracy by the Black-76 formula.

Recall that in electricity markets, forwards deliver over a period rather than at a fixed time T. A way of modelling such forwards is to introduce a dynamics $f(t, T^*)$, where $T^* \in (T_1, T_2)$ is some time in the delivery period $[T_1, T_2]$. A natural choice of T^* could be the middle point $T^* = (T_1 + T_2)/2$. It is a well-known empirical fact that in electricity markets forward prices do not converge to the spot prices if one approaches delivery time. Choosing T^* as the mid point of the delivery period will take this into account (see Barndorff-Nielsen et al. [8]). In the electricity markets, many options have exercise time equal to the beginning of delivery T_1 of the underlying forward, e.g. $\tau = T_1$. If the delivery period is relatively long, we will have that τ is relatively far from T^*. Hence, for a reasonably strong mean reversion β of the spikes, options on forwards in electricity markets can be priced with a high degree of accuracy by the Black-76 formula.

We illustrate our results with a numerical example. At $t = 0$, let the exercise time of the option be in $\tau = 10$ days and consider forwards with delivery times in $T = 10, ..., 40$ days. Assume that the speed of mean reversion is $\beta = 0.3466$, which corresponds to a half life of two days. Such a mean reversion rate is not unreasonable for spikes in electricity markets (see e.g. Benth et al. [15]). We model directly under the pricing measure Q. Let L be a compound Poisson process that has an exponential jump size distribution with mean equal 0.5 and a jump intensity of 5 jumps per month. This is a rather high number of spikes, but could mimic the situation in winter months, say, in the Nordic electricity market NordPool. The volatility the Brownian motion is $\sigma = 0.0158$ which corresponds to 30% annually. For simplicity, we assume that there is no seasonality, that is $\Lambda(t) = 1$ and fix the initial value of the forward to $x = 100$. We look at options at the money and assume $K = 100$. Using the Black-76 formula in Prop. 4.3.1, we get $C_{B76} = 1.9931$. We evaluate the option price $C(0, 10, T, 100)$ as in Prop. 4.3.2 with Monte Carlo simulation. For this purpose, the stochastic integral in Prop. 4.3.2 is discretized with a simple Euler scheme on a daily time grid. The price differences as well as the logarithmic price differences are plotted in Figure 4.1 together with the corresponding error bound from Prop. 4.3.2 (solid lines). The exponential decay of the error is in line with our theoretical results. If we consider an electricity forward with a monthly delivery period of 30 days, that is, $T_2 - T_1 = 30$, which starts at $T_1 = 10$, this would correspond to a $T = 25$ days if we let $T^* = (T_1 + T_2)/2$. Looking at Figure 4.1 we see that at this time, the prices using Prop. 4.3.2 and Prop. 4.3.1 are already very close. In fact, we have that $C(0, 10, 25, 100) = 1.9337$, implying that Black-76 is miss-pricing by only 3 %.

4.3.1 A transformed-based option pricing formula

For the sake of completeness, we include here a pricing formula for $C(t, \tau, T, x)$ based on the Fourier transform and the characteristic function of $f(t, T)$.

Recall the price $C(t, \tau, T, x)$ in Prop. 4.3.2. Denote the first expectation by I_1 and the second by I_2. In Lemma 4.3.6 we changed probability from Q to $\widetilde{Q}$ to reach the expression

$$I_1 = \mathbb{E}_{\widetilde{Q}}\left[\Phi\left(d_1\left(x, \int_t^\tau \mathrm{e}^{-\beta(T-s)}\, dL(s)\right)\right)\right], \tag{4.3.1}$$

where the logarithmic cumulant function $\phi_{\beta,\widetilde{Q}}(\theta)$ of $\int_t^\tau \exp(-\beta(T-s))\, dL(s)$ with respect to $\widetilde{Q}$ is

$$\phi_{\beta,\widetilde{Q}}(\theta) = \int_t^\tau \{\phi((1+\theta)\mathrm{e}^{-\beta(T-s)}) - \phi(\mathrm{e}^{-\beta(T-s)})\}\, ds\,. \tag{4.3.2}$$

We now want to express the expectations I_1 and I_2 by Fourier transforms.

Letting $d(x, v)$ be a generic notation for $d_1(x, v)$ and $d_2(x, v)$, we find that $\Phi(d(x, v)) \to 1$ when $v \to \infty$ since $d(x, v) \to \infty$ when $v \to \infty$. On the other hand, as $d(x, v) \to -\infty$ when $v \to -\infty$, we find $\Phi(d(x, v)) \to 0$. Hence, the function $v \mapsto \Phi(d(x, v))$ is not in $L^1(\mathbb{R})$. However, by damping it using an exponential function, we get an expression which is integrable:

Lemma 4.3.9. *For any $\alpha > 0$, the function $v \mapsto \exp(-\alpha v)\Phi(d(x, v))$ is integrable on $\mathbb{R}$. Here $d(x, v)$ is generic for $d_i(x, v)$, $i = 1, 2$.*

Proof. Since $0 \leq \Phi(y) \leq 1$, we have by Tonelli's theorem (see Folland [38])

$$\begin{aligned}
\int_{\mathbb{R}} \mathrm{e}^{-\alpha v}\Phi(d(x, v))\, dv &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-\alpha v} \int_{-\infty}^{d(x,v)} \mathrm{e}^{-y^2/2}\, dy\, dv \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} \mathrm{e}^{-\alpha v} \mathbf{1}_{y \leq d(x,v)}\, dv\, \mathrm{e}^{-y^2/2}\, dy \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \int_{\sigma\sqrt{\tau-t}(y-d(x,0))}^{\infty} \mathrm{e}^{-\alpha v}\, dv\, \mathrm{e}^{-y^2/2}\, dy \\
&= \frac{1}{\alpha\sqrt{2\pi}} \int_{\mathbb{R}} \mathrm{e}^{-\frac{1}{2}y^2 - \alpha\sigma\sqrt{\tau-t}(y-d(x,0))}\, dy < \infty\,.
\end{aligned}$$

Hence, the Lemma follows. □

In the next Lemma we compute the Fourier transform of the function $v \mapsto \Phi_\alpha(d(x,v)) := \exp(-\alpha v)\Phi(d(x,v))$:

Lemma 4.3.10. *The Fourier transform of $v \mapsto \Phi_\alpha(d(x,v))$ is*

$$\widehat{\Phi}_\alpha(y) = \frac{1}{\alpha + \mathrm{i}y} \exp\left(\frac{1}{2}(\alpha + \mathrm{i}y)^2\sigma^2(\tau - t) + (\alpha + \mathrm{i}y)d(x,0)\sigma\sqrt{\tau - t}\right).$$

Moreover, $\widehat{\Phi}_\alpha \in L^1(\mathbb{R})$.

Proof. By definition of the Fourier transform and the Fubini-Tonelli theorem (see Folland [38]), we find

$$\begin{aligned}
\widehat{\Phi}_\alpha(y) &= \int_{\mathbb{R}} \mathrm{e}^{-\alpha v}\Phi(d(x,v))\mathrm{e}^{-\mathrm{i}vy}\,dv \\
&= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\int_{-\infty}^{d(x,v)} \mathrm{e}^{-z^2/2}\,dz\,\mathrm{e}^{-(\alpha+\mathrm{i}y)v}\,dv \\
&= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\int_{\mathbb{R}} \mathbf{1}_{(z\le d_2(x,v))}\mathrm{e}^{-(\alpha+\mathrm{i}y)v}\,dv\,\mathrm{e}^{-z^2/2}\,dz \\
&= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\int_{\sigma\sqrt{\tau-t}(z-d_2(x,0))}^{\infty} \mathrm{e}^{-(\alpha+\mathrm{i}y)v}\,dv\,\mathrm{e}^{-z^2/2}\,dz \\
&= \frac{1}{\sqrt{2\pi}}\frac{1}{\alpha+\mathrm{i}y}\int_{\mathbb{R}} \mathrm{e}^{-(\alpha+\mathrm{i}y)\sigma\sqrt{\tau-t}z}\mathrm{e}^{-z^2/2}\,dz\,\mathrm{e}^{d_2(x,0)\sigma\sqrt{\tau-t}(\alpha+\mathrm{i}y)} \\
&= \frac{1}{\alpha+\mathrm{i}y}\mathbb{E}\left[\mathrm{e}^{-(\alpha+\mathrm{i}y)\sigma\sqrt{\tau-t}Z}\right]\mathrm{e}^{d_2(x,0)\sigma\sqrt{\tau-t}(\alpha+\mathrm{i}y)}.
\end{aligned}$$

Here, Z is a standard normally distributed random variable. This means that

$$\mathbb{E}\left[\mathrm{e}^{-(\alpha+\mathrm{i}y)\sigma\sqrt{\tau-t}Z}\right] = \mathrm{e}^{\frac{1}{2}(\alpha+\mathrm{i}y)^2\sigma^2(\tau-t)}.$$

This shows the Fourier transform of $\Phi_\alpha(x,v)$.

By taking absolute values, we find

$$|\widehat{\Phi}_\alpha(y)| = \frac{c}{\alpha^2+y^2}\mathrm{e}^{-\frac{1}{2}y^2\sigma^2(\tau-t)},$$

for a constant c independent of y. This shows that $\widehat{\Phi}_\alpha$ is an integrable function on $\mathbb{R}$. The proof is complete. □

Appealing to the inverse Fourier transform (see Folland [38]), we have the relation

$$\Phi(d(x,v)) = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\Phi}_\alpha(y)\mathrm{e}^{(\alpha+\mathrm{i}y)v}\,dy. \tag{4.3.3}$$

We apply this in order to express the call option price in terms of the Fourier transform of Φ_α and the characteristic function of L:

Proposition 4.3.11. *The call option price $C(t,\tau,T,x)$ in Prop. 4.3.2 can be expressed as*

$$C(t,\tau,T,x) = \frac{x}{2\pi}\int_{\mathbb{R}} \widehat{\Phi}_{1,\alpha}(y)\exp\left(\phi_{\beta,\widetilde{Q}}(\alpha+\mathrm{i}y)\right)\,dy - \frac{K}{2\pi}\int_{\mathbb{R}} \widehat{\Phi}_{2,\alpha}(y)\exp\left(\int_t^\tau \phi\left((\alpha+\mathrm{i}y)\mathrm{e}^{-\beta(T-s)}\right)\,ds\right)\,dy$$

for any $0<\alpha\leq 2$. We have introduced the notation $\widehat{\Phi}_{i,\alpha}$ to indicate that we use $d_i(x,v)$, $i=1,2$ as the function $d(x,v)$.

Proof. Using (4.3.3) it holds

$$\begin{aligned}\mathbb{E}&\left[\Phi\left(d_2\left(x,\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)\right)\right)\right]\\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\Phi}_{2,\alpha}(y)\mathbb{E}\left[\exp\left((\alpha+\mathrm{i}y)\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)\right)\right]\,dy\\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\Phi}_{2,\alpha}(y)\exp\left(\int_t^\tau \phi\left((\alpha+\mathrm{i}y)\mathrm{e}^{-\beta(T-s)}\right)\,ds\right)\,dy\,.\end{aligned}$$

Note that we must have $\alpha\leq 3$ in order for this to be well-defined, according to the exponential integrability condition (see Theorem 25.17(iii) in Sato [53]). This shows the second term in the price.

For the first term, we use the expectation under the probability $\widetilde{Q}$ as in (4.3.1). Again using the inverse Fourier transform along with Fubini-Tonelli's theorem,

$$\begin{aligned}\mathbb{E}_{\widetilde{Q}}&\left[\Phi\left(d_1\left(x,\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)\right)\right)\right]\\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\Phi}_{1,\alpha}(y)\mathbb{E}_{\widetilde{Q}}\left[\exp\left((\alpha+\mathrm{i}y)\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)\right)\right]\,dy\\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\Phi}_{1,\alpha}(y)\exp\left(\phi_{\beta,\widetilde{Q}}(\alpha+\mathrm{i}y)\right)\,dy\,.\end{aligned}$$

Note that $\phi_{\beta,\widetilde{Q}}(\alpha+\mathrm{i}y)$ is well-defined as long as $\alpha\leq 2$. This proves the first term, and the Proposition follows. □

We remark that the transformed-based pricing equation in the Proposition above lends itself to fast Fourier transform methods for numerical evaluation (see Eberlein et al. [34]).

4.4 Quadratic hedging of call options on forwards

We next consider hedging of the call option. Our market is incomplete, since the forward price dynamics is a jump-diffusion process. In this case there exists no self-financing portfolio in the underlying forward contract and a bank account replicating the option exactly. Instead, one must apply hedging strategies which minimizes, under some criterion, the hedging error. The hedging error is defined to be the difference between the terminal value of the hedging portfolio and the option, and we shall here look at hedges which minimize the variance, also called quadratic hedging. We refer the reader to Cont and Tankov [31] for a detailed discussion on incomplete markets and hedging, in particular quadratic hedging.

The next Proposition states the quadratic hedge position in the forward:

Proposition 4.4.1. *The quadratic hedge position $\psi(t)$ at time $t \leq \tau$ in the forward is given by*

$$\begin{aligned}\psi(t) = &\frac{\sigma^2}{\sigma^2 + \int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)^2 \,\ell(dz)} C_x(t,\tau,T,f(t,T)) \\ &+ \frac{\int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)\left(C(t,\tau,T,f(t,T)\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}}) - C(t,\tau,T,f(t,T))\right)\ell(dz)}{f(t,T)(\sigma^2 + \int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)^2 \,\ell(dz))}\,.\end{aligned}$$

Proof. First, let $\widetilde{C}(t,\tau,T,x) = \mathrm{e}^{-rt}C(t,\tau,T,x)$, the discounted option price. We know that the process $t \mapsto \widetilde{C}(t,\tau,T,x)$ is a martingale by the no-arbitrage pricing theory. Applying Itô's Formula for jump-diffusion, shows that

$$\begin{aligned}&d\widetilde{C}(t,\tau,T,f(t,T)) \\ &\quad= \sigma f(t,T)\widetilde{C}_x(t,\tau,T,f(t,T))\,dW(t) \\ &\qquad+ \int_{\mathbb{R}} \left\{\widetilde{C}(t,\tau,T,f(t,T)\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}}) - \widetilde{C}(t,\tau,T,f(t,T))\right\}\widetilde{N}(dt,dz)\,.\end{aligned}$$

If we let $\widetilde{V}(t) = \mathrm{e}^{-rt}V(t)$ be the discounted value of a self-financing portfolio (which is a martingale as well), then

$$d\widetilde{V}(t) = \psi(t)\mathrm{e}^{-rt}\,df(t,T) = \psi(t)\mathrm{e}^{-rt}f(t-,T)\left\{\sigma\,dW(t) + \int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)\,\widetilde{N}(dt,dz)\right\}.$$

Suppose that $V(0) = \widetilde{V}(0) = C(0,\tau,T,f(0,T))$. The hedging error is

$$\epsilon(\psi) = \widetilde{V}(\tau,\tau,T,f(\tau,T)) - \widetilde{C}(\tau,\tau,T,f(\tau,T))\,,$$

for $\tau \leq T$. But then it follows by Itô isometry for stochastic integration (see Cont and Tankov [31])

$$\begin{aligned}\mathbb{E}[\epsilon^2(\psi)] = &\int_0^\tau \mathbb{E}\left[(\widetilde{C}_x(s,\tau,T,f(s,T)) - \psi(s))^2\mathrm{e}^{-2rs}f^2(s,T)\right]\sigma^2\,ds \\ &+ \int_0^\tau\int_{\mathbb{R}} \mathbb{E}\Big[\Big(\widetilde{C}(s,\tau,T,f(s,T)\mathrm{e}^{z\mathrm{e}^{-\beta(T-s)}}) - \widetilde{C}(s,\tau,T,f(s,\tau,T,f(s,T)) \\ &-\psi(s)\mathrm{e}^{-rs}f(s,T)(\mathrm{e}^{z\mathrm{e}^{-\beta(T-s)}} - 1)\Big)^2\Big]\,\ell(dz)\,ds \\ = &\int_0^\tau \mathrm{e}^{-2rs}\mathbb{E}\left[f^2(s,T)\left\{\sigma^2\left(C_x(s,\tau,T,f(s,T)) - \psi(s)\right)^2 + \right.\right. \\ &+ \int_{\mathbb{R}}\left(\frac{C(s,\tau,T,f(s,T)\mathrm{e}^{z\mathrm{e}^{-\beta(T-s)}}) - C(s,\tau,T,f(s,T))}{f(s,T)}\right.\end{aligned}$$

$$-\psi(s)(\mathrm{e}^{z\mathrm{e}^{-\beta(T-s)}}-1)\Big)^2\,\ell(dz)\Big\}\Big]\,ds\,.$$

The first order condition for the minimizer of this quadratic expression solves

$$\begin{aligned}
&\psi(t)\left(\sigma^2+\int_{\mathbb{R}}\left(\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}}-1\right)^2\,\ell(dz)\right)\\
&\quad=\sigma^2 C_x(t,\tau,T,f(t,T))\\
&\qquad+f^{-1}(t,T)\int_{\mathbb{R}}(\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}}-1)\\
&\qquad\qquad\times\left(C(t,\tau,T,f(t,T)\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}})-C(t,\tau,T,f(t,T))\right)\,\ell(dz)
\end{aligned}$$

Hence, the Proposition follows. □

Our goal now is to show that this quadratic hedge converges to the delta hedging strategy $C_x(t,\tau,T,x)$ of the Black-76 call. To have a more suggestive notation, we let

$$C(t,x;\beta):=C(t,\tau,T,x)$$

and we recall from theorem 4.3.8 that

$$\lim_{\beta\downarrow 0}C(t,x;\beta)=C_{\mathrm{B76}}(t,x)\,,$$

with the obvious meaning of the short-hand notation $C_{\mathrm{B76}}(t,x)$. The Black-76 formula is the price of a call in a complete market, and the hedge position in the forward is given by the derivative of the call price with respect to the forward, $C_{\mathrm{B76},x}(t,x)$ (see Cont and Tankov [31], say). We show, in a sequence of Lemmas, that

$$\psi(t)\to C_{\mathrm{B76},x}(t,x)\,,$$

when $T-t\to\infty$. Moreover, we show that the convergence is uniform with an exponential rate given by β.

First, we recall the delta hedge in the Black-76 market:

Proposition 4.4.2. *The delta hedge of Black-76 is*

$$C_{B76,x}(t,x)=\Phi(d_1(x))\,,$$

with $d_1(x)$ defined in Prop. 4.3.1.

Proof. This is a straightforward application of the result in Prop. 4.3.1. □

The next Proposition shows that the derivative of $C(t,x;\beta)$ has the same shape as the delta hedge in the Black-76 market:

Proposition 4.4.3. *For every $t \leq \tau \leq T$ and $\beta > 0$ it holds*

$$
\begin{aligned}
C_x(t,x;\beta) = \mathbb{E}\Bigg[\exp\left(\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s) - \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds\right) \\
\times\, \Phi\left(d_1\left(x, \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,dL(s)\right)\right)\Bigg]\,,
\end{aligned}
$$

where $d_1(x,v)$ is defined in Prop. 4.3.2.

Proof. A direct derivation of the expression in Prop. 4.3.2 yields,

$$
\begin{aligned}
C_x(t,x;\beta) &= \mathbb{E}\left[\mathrm{e}^{Z-\int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds}\Phi\left(d_1(x,Z)\right)\right] \\
&\quad + x\mathbb{E}\left[\mathrm{e}^{Z-\int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds}\Phi'\left(d_1(x,Z)\right)\frac{\partial d_1(x,Z)}{\partial x}\right] \\
&\quad - K\mathbb{E}\left[\Phi'\left(d_2(x,Z)\right)\frac{\partial d_2(x,Z)}{\partial x}\right]\,,
\end{aligned}
$$

with $Z = \int_t^\tau \exp(-\beta(T-s))\,dL(s)$. We focus on the last two terms, which we show are adding up to zero. From the definitions of $d_i(x,v)$, $i = 1,2$,

$$
\begin{aligned}
\frac{\partial d_1(x,v)}{\partial x} &= \frac{\partial d_2(x,v)}{\partial x}\,, \\
\frac{\partial d_1(x,v)}{\partial x} &= \frac{1}{x}\frac{1}{\sigma\sqrt{\tau-t}}\,.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
& x\mathbb{E}\left[\mathrm{e}^{Z-\int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds}\Phi'\left(d_1(x,Z)\right)\frac{1}{x\sigma\sqrt{\tau-t}}\right] - K\mathbb{E}\left[\Phi'\left(d_2(x,Z)\right)\frac{1}{x\sigma\sqrt{\tau-t}}\right] \\
&\qquad = \frac{1}{\sigma\sqrt{\tau-t}}\mathbb{E}\left[\mathrm{e}^{Z-\int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds}\Phi'\left(d_1(x,Z)\right) - \frac{K}{x}\Phi'\left(d_2(x,Z)\right)\right]\,.
\end{aligned}
$$

As Φ is the cumulative distribution function of a standard normal variable, we find from the definition of $d_1(x,v)$ that

$$
\Phi'\left(d_2(x,Z)\right) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}d_2^2(x,Z)\right) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(d_1(x,Z)+\sigma\sqrt{\tau-t})^2\right)\,.
$$

But since

$$
d_1(x,Z)\sigma\sqrt{\tau-t} = \ln x - \ln K + Z - \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds - \frac{1}{2}\sigma^2(\tau-t)\,,
$$

we have

$$
\mathrm{e}^{Z-\int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds}\Phi'\left(d_1(x,Z)\right) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}d_2^2(x,Z)\right)\frac{K}{x}\,.
$$

This concludes the proof of the Proposition. □

In the first lemma, we study the convergence of the "variance" term from the jumps:

Lemma 4.4.4. *For $t \leq T$, it holds*

$$\int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)^2 \,\ell(dz) \leq c\mathrm{e}^{-2\beta(T-t)}\,,$$

for a constant $c > 0$.

Proof. Although we have implicitly estimated this convergence in the Lemmas of the previous Section, we spell it out here for the convenience of the reader. For any positive constant $k \leq 1$ we have

$$|\mathrm{e}^{kz} - 1| \leq k|z| \sum_{n=1}^{\infty} \frac{k^{n-1}|z|^{n-1}}{n!} \leq k|z| \sum_{n=0}^{\infty} \frac{|z|^n}{n!} = k|z|\mathrm{e}^{|z|}\,.$$

But then

$$\begin{aligned}\int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)^2 \,\ell(dz) &\leq \int_{\mathbb{R}} |z|^2\mathrm{e}^{2|z|}\,\ell(dz)\mathrm{e}^{-2\beta(T-t)} \\ &\leq \left\{\int_{|z|\leq 1} z^2\,\ell(dz)\mathrm{e}^2 + \int_{|z|>1} z^2\mathrm{e}^{2|z|}\,\ell(dz)\right\}\mathrm{e}^{-2\beta(T-t)}\,.\end{aligned}$$

The Lemma follows from the exponential moment condition on $\ell(dz)$ and the condition that ℓ is a Lévy measure. □

A convenient property of the option price is that it is uniformly Lipschitz, as the next Lemma shows.

Lemma 4.4.5. *For every $t \leq \tau \leq t$ and $\beta > 0$, we have*

$$|C(t, x; \beta) - C(t, y; \beta)| \leq |x - y|\,,$$

for all $x, y \geq 0$.

Proof. By the mean-value theorem we find

$$|C(t, x; \beta) - C(t, y; \beta)| = |C_x(t, z; \beta)||x - y|\,,$$

for some $z \geq 0$. From Prop. 4.4.3 it follows

$$\begin{aligned}|C_x(t, z; \beta)| &= \mathbb{E}\left[\exp\left(\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s) - \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds\right)\right. \\ &\qquad \left.\times\,\Phi\left(d_1\left(z, \int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s)\right)\right)\right] \\ &\leq \mathbb{E}\left[\exp\left(\int_t^\tau \mathrm{e}^{-\beta(T-s)}\,dL(s) - \int_t^\tau \phi(\mathrm{e}^{-\beta(T-s)})\,ds\right)\right] \\ &= 1\,.\end{aligned}$$

The Lemma follows. □

We present our convergence result on the quadratic hedge in the next Theorem:

Theorem 4.4.6. *For $t \leq \tau \leq T$ it holds that*

$$\sup_{x \geq 0} |\psi(t) - C_{B76,x}(t,x)| \leq c\mathrm{e}^{-\beta(T-\tau)},$$

for a positive constant c.

Proof. By the triangle inequality it holds

$$\begin{aligned}
|\psi(t) - C_{\mathrm{B76},x}(t,x)| \leq & \left| \frac{\sigma^2}{\sigma^2 + \int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)^2 \,\ell(dz)} C_x(t,x;\beta) - C_{\mathrm{B76},x}(t,x) \right| \\
& + \frac{1}{\sigma^2 + \int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)^2 \,\ell(dz)} \\
& \times \int_{\mathbb{R}} |\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1| \left| \frac{C(t, x\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}};\beta) - C(t,x;\beta)}{x} \right| \ell(dz) \\
\leq & \, |C_x(t,x;\beta) - C_{\mathrm{B76},x}(t,x)| \\
& + \frac{\int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)^2 \,\ell(dz)}{\sigma^2 + \int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)^2 \,\ell(dz)} C_{\mathrm{B76},x}(t,x) \\
& + \frac{\int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)^2 \,\ell(dz)}{\sigma^2 + \int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)^2 \,\ell(dz)} .
\end{aligned}$$

In the last inequality we applied the Lipschitz continuity of C in Lemma 4.4.5. But from Prop. 4.4.2 we have that $C_{\mathrm{B76},x}(t,x) \leq 1$. Moreover,

$$\frac{\int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)^2 \,\ell(dz)}{\sigma^2 + \int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)^2 \,\ell(dz)} \leq \frac{1}{\sigma^2} \int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)^2 \,\ell(dz) .$$

This implies that

$$|\psi(t) - C_{\mathrm{B76},x}(t,x)| \leq |C_x(t,x;\beta) - C_{\mathrm{B76},x}(t,x)| + \frac{2}{\sigma^2} \int_{\mathbb{R}} (\mathrm{e}^{z\mathrm{e}^{-\beta(T-t)}} - 1)^2 \,\ell(dz) .$$

Invoking Prop. 4.4.3, and using Lemmas 4.3.6-4.3.7 the first term on the right hand side can be bounded uniformly in x by $\exp(-\beta(T-\tau))$. Hence, we conclude the result by appealing to Lemma 4.4.4. □

Not surprisingly, the convergence rate of the delta hedge is equal to the one for the prices. Thus, when $\beta(T-\tau)$ is sufficiently big, the quadratic hedge of the call option will be approximately equal to the Black-76 delta hedge. Again, referring back to electricity forwards, we may have this situation when the delivery period $[T_1, T_2]$ is relatively long compared with the speed of mean reversion β, letting $T = (T_1 + T_2)/2$.

4.5 Conclusion

Based on a generalization of the popular two-factor spot price model of Gibson and Schwartz [42], we show that call options written on forwards can be approximated by the Black-76 price in many situations. The logarithmic spot price dynamics consists of a non-stationary drifted Brownian motion factor, and a stationary factor modelled as a Lévy-driven Ornstein-Uhlenbeck process. The forward price becomes reasonably analytic under this spot model, and we derive the price of call options based on Fourier methods.

It is demonstrated that the option prices converge exponentially to the Black-76 price in terms of the speed of mean-reversion of the stationary factor in the spot price and the time left to maturity of the forward from the exercise time of the call. In many power markets, the stationary factor has a rather high speed of mean-reversion as this is modelling the spiky behaviour of spot prices. For options with exercise time relatively far from the delivery time of the forwards, the price will therefore be approximately given by the Black-76 formula. On the other hand, if the difference between time of delivery and exercise of option is small, the option price may be significantly far away from Black-76, unless the speed of mean-reversion is huge.

Typically, in gas and electricity, forwards deliver over a specified period like a month. In our framework we suggest to take delivery period forwards into account by assuming their dynamics being given by a forward delivering in the middle of the delivery period. Combining this approach with a typically high speed of mean-reversion, we can conclude that call options on electricity and gas forwards may be priced reasonably accurately by the Black-76 formula. In other words, we may completely ignore the spikes and non-Gaussian effects in the pricing, as these are ”killed” by the delivery period of the forward. A numerical example further argue for this.

As our model for the spot and forward prices leads to an incomplete market, we cannot hedge the call option. However, the quadratic hedging strategy minimizing the L^2-distance between the call payoff and a portfolio in the underlying forward can be derived in terms of the option price. It is shown that the quadratic hedge can be approximated by the delta-hedge from Black-76. Not surprisingly, the hedge tends exponentially to the Black-76 delta hedge at the same rate as the option price.

There exist several interesting extensions of our results that could be worthwhile pursuing. For example, empirical studies of spot prices of electricity suggest that the stationary factor can be better modelled using a more general continuous time autoregressive moving average dynamics than the ”AR(1)” Ornstein-Uhlenbeck process (see Garcia, Klüppelberg and Müller [39]). Another extension is to let the non-stationary factor be non-Gaussian, which is relevant in electricity (see e.g. Benth et al. [16]). Of course, such a spot model would not yield a convergence of option prices to the Black-76 formula as this rests on the Brownian motion driving the non-stationary part. A completely different path to follow is to check different hedging strategies than the quadratic one to analyse a possible convergence to the delta hedge of Black-76. This would lead into a different set-up for pricing and hedging of the options.

From an empirical point of view, it would be interesting to check our results with real option data in various markets. An immediate problem with such a study is that the liquidity in many energy option markets is rather low. Also, as we have mentioned in the above paragraphs, more sophisticated spot and forward models may be needed to reach firm conclusions. In any case, our analysis points towards the fact that the non-stationary factor is decisive for pricing and hedging options on forwards in energy markets.

CHAPTER 5

Article 4: "Pricing Futures and Options in Electricity Markets"

Fred Espen Benth and Maren Diane Schmeck

Abstract

In this paper we derive power futures prices from a two-factor spot model being a generalization of the classical Schwartz-Smith commodity dynamics. We include non-Gaussian effects by introducing Lévy processes as the stochastic drivers, and estimate the model to data observed at the European Electricity Exchange in Germany. The spot and futures price models are fitted jointly, including the market price of risk parameterized from an Esscher transform. We apply this model to price call and put options on power futures. It is argued theoretically that the pricing measure for options may be different to the pricing measure of futures from spot in power markets due to the non-storability of the electricity spot. Empirical evidence pointing to this fact is found from option prices observed at the European Electricity Exchange.

5.1 Introduction

In the last two decades markets for power have been liberalized in Europe and other places world-wide. Nowadays, we find well-functioning markets for purchase of electricity in many countries on the European continent, in the Nordic countries and in UK. Furthermore, there exists markets in North America, Australia and some places in Asia. Typically, these markets separate between a day-ahead spot market for electricity, and financial contracts for future delivery of power. In some, more developed markets, one also trades in derivatives like plain vanilla call and put options on the futures contracts. This takes place in for example the Nordic market NordPool and the German market European Electricity Exchange (EEX).

In this paper we focus the attention on pricing spot, futures and options jointly in the power market. Our aim is to argue for a separation of the modelling of the risk premium charged in the futures market and the risk neutral measure used for options pricing. The classical approach to futures pricing is to specify a stochastic dynamics of the spot price, and define the futures price as the conditional risk-adjusted expected average spot price over the delivery period. The risk-adjustment is modelled by a specification of pricing probabilities, which changes the characteristics of the spot dynamics (see Benth *et al.* [19] for a discussion and application of this approach to energy markets). Usually, as this approach yields a risk neutral (or martingale) dynamics of the futures price, one would price options using the same probability. We argue here that there is no violation of no-arbitrage pricing to have another pricing measure for options, as long as this is an equivalent martingale measure for the futures price dynamics. The economic argument in favour of this is the non-storability of the electricity spot price.

Based on a small data set of option prices at the EEX, we also argue empirically for this possibility. Fitting a two-factor model for the spot price dynamics to EEX data, we price futures and calibrate the risk premium using a parametric class of pricing probabilities stemming from the Esscher transform (see Benth *et al.* [19]. Although the access to option data at the EEX is poor due to a rather illiquid market, we find evidence for a different risk neutral pricing measure than the one used to derive futures prices from the spot dynamics. We benchmark our results to the Black-76 prices derived from historical volatility.

Our two-factor spot model is a generalization of the Schwartz-Smith dynamics (see Schwartz and Smith [55]), consisting of a long-term non-stationary factor and a short-term stationary factor. The Schwartz-Smith model has been applied to electricity markets by Lucia and Schwartz [47], who analysed spot and futures data at the NordPool market. As the Schwartz-Smith model is Gaussian, it fails to account for the large spikes in the market. We extend the model to include Lévy process driven noises, which also accounts for the high variability in EEX prices in non-spike periods. Our proposed model is a simplification of the dynamics proposed and analysed in Benth *et al.* [16] and Barndorff-Nielsen *et al.* [7]. The fitting of the spot and futures dynamics goes by filtering the non-stationary factor by using futures prices from contracts far from delivery.

The presentation of our results are separated into several sections. In the next section we present the rationale behind pricing of futures in power markets. Furthermore, we discuss the pricing of options, and why one may use a different probability for this purpose. Section 3 first defines the two-factor spot model, and presents theoretical futures prices based on this dynamics. The joint spot and futures price model is estimated to EEX data in section 4, while section 5 analyses empirically the option pricing performance of our futures price model. This section argues in favour of a different pricing measure for options. Finally, in section 6, we conclude and outline some future research directions.

5.2 The relation between spot, futures and options in power markets

Typically, the liberalized power markets are divided into a day-ahead spot market, a financial market for futures (and/or forwards[1]) contracts on power, and a market for plain vanilla call and put options on the futures. The forward contracts deliver the underlying power over an agreed period of time, and the delivery is settled financially, that is, the money-equivalent of

[1]Some markets have both forwards and futures traded. We shall not make a distinction between these two asset classes here, but stick to the notion of futures.

the spot is delivered. These contracts are denominated in a "currency" per MWh and work essentially as a swap contract where one exchanges a floating spot price against a fixed over the contracted period.

For example, in the German EEX market the swaps have delivery periods being months, quarters or years. The swap price is naturally denominated in Euro per MWh, and the contract is accounted against the hourly power spot price. One distinguishes between base and peak load contracts, where the peak load take into account only the power spot prices in the peak hours, defined as the working days from 8 in the morning to 8 in the evening. The base load contracts are settled against the spot price of all hours in the delivery period.

The power spot prices are determined in an auction-based system, where the traders hand in prices and volumes for production or consumption for given hours the next day. Based on these bids, the exchange creates demand and supply curves for each hour the following day, and at 2pm the EEX publishes these spot prices for the 24 hours next day. We emphasise that the trade in the power spot market is physical, and one therefore needs to have facilities for either producing or consuming (retailing) electricity. Unlike most other assets that can be traded, one cannot form a portfolio and use the spot for investment or speculation purposes. By the very nature of electricity, it is not possible to store. There are some exceptions, since one may in fact use water reservoirs, say, as storage of power in terms of potential energy. However, this is only possible for a limited segment of the market, namely the hydro power producers.

The options traded in the market are written on specific financial swap contracts. At the EEX power options are written on the Phelix Base futures with monthly, quarterly and yearly delivery periods. The EEX offers only European style call and put options, where the exercise takes place four trading days prior to the beginning of the delivery period of the underlying futures.

Let us discuss at a more technical level the relationship between spot prices, swaps, and options. For illustration, consider first a market where the spot is a liquidly tradeable asset, like for example an exchange-traded stock. We denote $S(t)$ as the spot price at time $t \geq 0$, and consider a forward contract which delivers the spot at a maturity time T. The forward price at time $t \leq T$ is denoted by $f(t, T)$, and from standard no-arbitrage arguments based on the cash and carry strategy (see *e.g.* Duffie [33]), it can be determined as

$$f(t, T) = S(t)\mathrm{e}^{r(T-t)} . \tag{5.2.1}$$

Here, $r > 0$ is the deterministic risk-free interest rate, where we have supposed that interest rates are continuously compounded. As is known from classical financial theory, (5.2.1) can be established without any model assumptions on the spot price.

Assume that we are given a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,\widetilde{T}]}, P)$. We interpret $\widetilde{T} < \infty$ as the time horizon of the market, including the maturities of all options and futures relevant in our analysis. If $S(t)$ is a semimartingale process, then there exists (at least one) equivalent martingale measure Q such that

$$f(t, T) = \mathbb{E}_Q[S(T) \,|\, \mathcal{F}_t] . \tag{5.2.2}$$

We refer to Shiryaev [56] for the rigorous argumentation with conditions leading to this representation of $f(t, T)$. In a complete market, that is, a market where all derivatives on S can be replicated, the probability measure Q is uniquely defined. In the case of an incomplete market, one may have many such measures Q. The question is to determine *one* relevant for pricing of derivatives. But, once such a measure is pinned down, we can price futures and next use the same probability for pricing options. Thus, for example the price of a European option with payoff $g(S(f(\tau,T))$ at exercise time $\tau \leq T$ becomes

$$C(t) = \mathrm{e}^{-r(\tau-t)}\mathbb{E}_Q[g(f(\tau,T)) \mid \mathcal{F}_t]\,,$$

for $0 \leq t \leq \tau$. Note that we use the same Q for both the forward and the option, as is the customary when pricing several derivatives based on an asset in an incomplete market situation. Note, however, that we may use different equivalent martingale measures for pricing different derivatives, as long as there exists at least *one* measure Q that is an equivalent martingale measure for *all* products.

To see this, suppose that we have two derivatives on the spot with payoffs given by the random variables X and Y, respectively. Let the prices at time zero be $C_X = \mathbb{E}_{Q_X}[X]$ and $C_Y = \mathbb{E}_{Q_Y}[Y]$, where we for the moment assume that the interest rate is zero to simplify the argument. The probabilities Q_X and Q_Y are equivalent martingale measures. If there exists an equivalent martingale measure Q, such that the price processes S, C_X and C_Y, are all Q-martingales, then the market is arbitrage-free. However, as long as Q is equivalent to P, it has to be equivalent to Q_X and Q_Y as well. Furthermore, by the no-arbitrage theory we must have that $C_X = \mathbb{E}_Q[X]$ and $C_Y = \mathbb{E}_Q[Y]$. This implies that

$$\mathbb{E}_{Q_X}\left[X\frac{dQ}{dQ_X}\right] = \mathbb{E}_{Q_X}[X]\,,$$

and

$$\mathbb{E}_{Q_Y}\left[Y\frac{dQ}{dQ_Y}\right] = \mathbb{E}_{Q_Y}[Y]\,.$$

These two equalities put strong conditions on the range of possible probabilities Q_X, Q_Y and Q.

In the case of power markets, the situation is completely different since the probability measure used to price futures can theoretically be completely detached from the measure pricing options on futures. As we have already argued, the power spot price cannot be traded in the normal financial sense, and it works as a *reference index* for the settlement of forward contracts. With this view at hand, the *pricing measure* Q used to derive the forward price on the spot does not need to be an *equivalent martingale* measure, but is required only to be an *equivalent* measure. However, the forward is a tradeable asset and its price dynamics must be a Q-martingale in order for the market to be free of arbitrage opportunities. Pricing using conditional expectation as in (5.2.2) ensures this by definition.

In a specification of the market, one would typically model the spot price evolution using some stochastic process $S(t)$, and choose a parametric class of equivalent probability measures Q. Based on a selected probability Q from this class, the standard approach to price electricity futures is to define it as

$$F(t,T_1,T_2) = \mathbb{E}_Q\left[\frac{1}{T_2-T_1}\int_{T_1}^{T_2} S(t)\,dt \mid \mathcal{F}_t\right]\,. \tag{5.2.3}$$

Here, we consider a contract delivering electricity over the time interval $[T_1,T_2]$, and the contract is entered at time $t \leq T_1$, with settlement at the end of the delivery period T_2. Note that the price is denoted in MWh, and therefore is normalized by the length of the delivery period. This gives a theoretical swap price dynamics which we next calibrate to the observed prices by fitting the parameters of the probabilities Q. This will pin down a probability $\widehat{Q}$ under which we model the risk-neutral forward price dynamics. Note that the risk-neutral dynamics of F is a $\widehat{Q}$-martingale. Since by construction $\widehat{Q}$ is equivalent to P, we can also (in principle) derive the market dynamics of the forward. Note that in general $\widehat{Q}$ is not a probability for which the spot price dynamics becomes a martingale after discounting.

In reality, the above procedure in specifying a probability $\widehat{Q}$ for pricing futures is an approach to find a parametric representation of the price process $F(t, T_1, T_2)$, where we calibrate to represent *the risk premium* in the market, that is, to explain the difference between the observed forward prices and the predicted average spot price. The latter is calculated by relation (5.2.3) using $Q = P$. Apriori there are two extreme choices one can make on Q. First, ignoring the existence of a risk premium, one could select $Q = P$. Alternatively, assuming the electricity spot is tradeable, one could force Q to be a *martingale* measure. Note that depending on the model for S, one could have many possible martingale measures, so the latter choice is not necessarily unique. Both alternatives are theoretically viable, but hardly reasonable from the characteristics of electricity markets.

Our next problem is to price call and put options written on the futures. Following the standard no-arbitrage pricing framework discussed above, a first thought would be to use $\widehat{Q}$ and compute the option price using this probability. To be more specific, let us suppose that we have a call option with exercise time $\tau \leq T_1$ written on a swap with dynamics $F(t, T_1, T_2)$ given in (5.2.3) for the pricing measure $\widehat{Q}$. The price of this call at time $t \leq \tau$ is be

$$C(t) = \mathrm{e}^{-r(\tau-t)}\mathbb{E}_{\widehat{Q}}\left[\max(F(\tau, T_1, T_2) - K, 0) \,|\, \mathcal{F}_t\right] .$$

However, in general, there will exist several equivalent measures Q for which $t \mapsto F(t, T_1, T_2)$ is a Q-martingale. In fact, since typically the power spot price dynamics involves jump processes, the forward price will also follow a jump dynamics as well. Under certain conditions, such models admit the existence of a continuum of equivalent martingale measures Q. In this case we pin down a pricing measure $\widetilde{Q}$ by selecting it from a parametric class of equivalent martingale measures Q for $F(t, T_1, T_2)$. One could derive this probability by calibrating to observed option prices in the market, or to appeal (partial) hedging arguments (see Cont and Tankov [31] for a discussion of hedging and pricing in incomplete markets).

Note that finding $\widetilde{Q}$ for option pricing follows in principle the same scheme as choosing $\widehat{Q}$ for the forward prices. The fundamental difference is that $\widehat{Q}$ does not need to be a martingale measure for the spot price, whereas $\widetilde{Q}$ has to be a martingale measure for the forward price. Both probability measures are equivalent to P. In the next sections we shall estimate a particular two-factor model to spot price data collected from the EEX, and apply this to forward pricing based on a class of probabilities defined by Esscher transformation. Using option price data, we shall argue that the spot-forward probability $\widehat{Q}$ is not the right probability for pricing options on the forward, pointing towards $\widetilde{Q} \neq \widehat{Q}$.

Our analysis is not restricted to power markets only. In the weather markets, like the temperature market at the Chicago Mercantile Exchange (CME), futures on temperature indices measured in various cities world-wide are traded. In addition, plain vanilla call and put options on these futures are traded. The underlying "spot" price here is the temperature in a given city, for example Chicago itself. Given a stochastic model for the temperature

$S(t)$, one can derive the resulting futures price written on an index of the temperature. Typically, one chooses to price using a conditional expectation analogous to (5.2.3), where a pricing measure is selected. Obviously, temperature itself is not a tradeable commodity, and we can use the same argumentation as above to defend choosing the pricing probabilities which are not necessarily martingale measures for the temperature dynamics. On the other hand, the futures contracts are tradeable financial assets, and to price the options with these as underlying, we need to use a probability measure Q which turns the futures price into a Q-martingale. As in the case of power, the futures pricing measure $\widehat{Q}$ does not need to be the same as the option pricing measure $\widetilde{Q}$. We note in passing that CME also organize a market for precipitation derivatives based on snow and rainfall indices in some cities in the US. Further, there has been trials to create an organized market for wind futures and options at the now closed US Futures Exchange. Here our discuss makes sense as well.

5.3 The spot price dynamics and implied forward prices

We consider a simple arithmetic two-factor spot price dynamics in the spirit of Lucia and Schwartz [47]. The occurrence of negative spikes at the EEX, and, even more, the observation that these spikes may even lead to negative prices, indicate that an arithmetic model may be suitable. To this end, suppose that $S(t)$ follows the dynamics

$$S(t) = \Lambda(t) + X(t) + Y(t)\,. \tag{5.3.1}$$

Here, $\Lambda : [0, \widetilde{T}] \mapsto \mathbb{R}$ is a measurable deterministic function, modelling the mean seasonal variation in spot prices. Usually, this function consists of a linear trend and a periodic function (a linear combination of sines and cosines, with different frequencies), and as such is a smooth function. The *base component* $X(t)$ in the spot price dynamics is assumed to be non-stationary and defined to be a Lévy process, i.e.,

$$dX(t) = dL_1(t)\,.$$

In Lucia and Schwartz [47], it is assumed that $L_1(t) = \gamma t + \sigma B(t)$ with γ and σ being constants and $B(t)$ a Brownian motion. The volatility σ is naturally assumed to be positive. One may think of the base component as stochastic variations from market activity as well as long term effects like inflation in fuel prices and limited resources, as well as entry of new sources of energy (like renewables). As it will turn out from our empirical analysis of EEX spot price data, a drifted Brownian motion is unsuitable for modelling the true dynamics of the non-stationary term, and a Lévy process is much more appropriate.

Typically in power markets spot prices may exhibit random shocks due to imbalances in supply and demand. These shocks are seen as spikes in the price path, imposed from an unexpected increase in demand due to colder weather, say, or shut down of a major power plant yielding a drop in supply. The prices will in both these cases exhibit a major price jump upward, which is followed typically by a strong decline since demand will be significantly reduced by higher prices, or expensive power production plants are ramped up (like coal-fired plants in Denmark in the NordPool area). In the EEX market one observes many negative spikes, which is caused by wind power mainly. By political legislation, wind power and other renewable energy sources have priority into the electricity grid, and hence an unexpected increase in wind power production (due to more wind where the farms are...) may create bigger than expected supply (since it takes time to ramp down or adjust other power plants fueled by gas and coal or producing nuclear energy). In fact, one observes negative prices in the EEX market due to over-supply, where some producers chooses to pay for power consumption rather than shut down their production.

From this discussion, we see that there is ample evidence for a mean-reverting short-time factor of the form

$$dY(t) = -\eta Y(t)\, dt + dL_2(t)\,.$$

Here, the constant $\eta > 0$ is expected to be rather big, since spikes created by the Lévy process $L_2(t)$ is reverting fastly back to normal price levels. We suppose that $L_2(t)$ may have both positive and negative jumps, that is, $L_2(1)$ is distributed on $\mathbb{R}$.

Notice that in Lucia and Schwartz [47], both an arithmetic and geometric two-factor model were analysed theoretically and empirically on NordPool data. In their approach, the second factor Y was also assumed to be driven by a Brownian motion. We believe that a jump factor for the noise is more appropriate in order to explain the sudden spikes in prices, exhibiting a jump like behaviour in the price path. Also, most empirical studies of power spot prices point strongly towards non-Gaussianity in prices, and hence the need to use other processes than the Brownian motion to drive the dynamics (see discussion in Benth *et al.* [19]). We remark that Lucia and Schwartz [47] let the short and long term factors correlate through their driving noise.

We denote $L = (L_1, L_2)$, and assume that L is a bivariate Lévy process with cumulant (log-characteristic function) defined by

$$\psi(\mathbf{x}) = \mathrm{i}\mu'\mathbf{x} - \frac{1}{2}\mathbf{x}'C\mathbf{x} + \int_{\mathbb{R}^2} \mathrm{e}^{\mathrm{i}\mathbf{x}'\mathbf{z}} - 1 - \mathrm{i}\mathbf{x}'\mathbf{z}1(|\mathbf{z}| \le 1)\, \ell(d\mathbf{z})\,,$$

with $\mathbf{x} = (x, y)' \in \mathbb{R}^2$, $\mu \in \mathbb{R}^2$, C a symmetric non-negative definite 2×2 matrix and $\ell(d\mathbf{z})$ a Lévy measure on $\mathbb{R}^2\backslash\{\mathbf{0}\}$. Here $\mathbf{x}'$ denotes the transpose of the vector, and $\mathrm{i} = \sqrt{-1}$ is the imaginary unit. In the case of independence between L_1 and L_2, we can express the cumulant as a sum

$$\psi(x, y) = \psi_1(x) + \psi_2(y)$$

where ψ_i, $i = 1, 2$ are cumulants of the univariate Lévy processes L_1 and L_2. Our general model allows for a dependency between L_1 and L_2, although we shall assume independence in the empirical study on EEX data below.

In Benth *et al.* [16] they use a more general model. The stationary short time variations are modelled as a continuous-time autoregressive moving average (CARMA) process, where the driving process L_2 is an α-stable Lévy process. As it includes mean reversion, a CARMA model is comparable to the standard approach of commodity spot price modelling, that is, to describe the spot as a sum of several Ornstein-Uhlenbeck processes with different speeds of mean reversion and stochastic drivers (see Benth *et al.* [19]). In Benth *et al.* [16], a CARMA(2,1) dynamics is proposed and fitted empirically to EEX spot price data. Such a dynamics is similar to a two-factor model, with each factor being an Ornstein-Uhlenbeck process. Although we find strong indications of a two-factor dynamics in our empirical study, we simplify the considerations here to a one-factor model as a first order approximation of the short-term factor. This makes the fitting of data significantly easier, and is in line with the more classical two-factor model of Lucia and Schwartz [47]. Moreover, it turns out that we can do well with a much more regular Lévy process that the α-stable to model the random fluctuations.

Our first concern is to introduce a parametric class of equivalent probabilities Q which is appropriate for pricing swaps. For $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$, define the equivalent probability Q_θ, where the density process of Q_θ with respect to P is

$$\left.\frac{dQ_\theta}{dP}\right|_{\mathcal{F}_t} = \exp\{\theta L(t) - \psi(-\mathrm{i}\theta)\, t\}\,.$$

In order for this to be well-defined, we must of course assume exponential integrability conditions on $L(1)$. Hence, suppose from now on that there exists a constant $c > 0$ such that

$$\int_{\mathbb{R}^2} \mathrm{e}^{\mathbf{x}'\mathbf{z}}\, \ell(d\mathbf{z}) < \infty\,, \tag{5.3.2}$$

for all $|\mathbf{x}| \leq c$. This ensures finite exponential moments for $L(1)$ up to order c.

The probability Q_θ parameterized by θ is known as the Esscher transform of L (see Benth *et al.* [19]). The probability Q_θ is equivalent to P by definition of the Radon-Nikodym densities. We emphasize, however, that we do not demand Q_θ to be a *martingale* measure, in the sense that the power spot dynamics becomes a Q_θ-martingale (the reader should note that this is technically impossible anyway with the Esscher transform on an Ornstein-Uhlenbeck process, see Benth and Sgarra [20]). The reason is the non-storability of the spot which makes it non-tradeable, that is, one cannot create portfolios with spot investments in electricity. Once purchased, it must be consumed. The parameter θ is restricted to the subspace of $\mathbb{R}^2$ defined by $|\theta| \leq c$.

In the next Lemma we characterize the process L under Q_θ:

Lemma 5.3.1. *The process L is a Lévy process with respect to Q_s with cumulant function*

$$\psi_{Q_\theta}(\mathbf{x}) = \psi(\mathbf{x} - \mathrm{i}\theta) - \psi(-\mathrm{i}\theta)$$

Hence, the drift is

$$\mu + \theta' C + \int_{|\mathbf{z}|<1} (\mathrm{e}^{\theta\mathbf{z}} - 1)\mathbf{z}\, \ell(d\mathbf{z})$$

and Lévy measure

$$\ell_{Q_\theta}(d\mathbf{z}) = \mathrm{e}^{\theta\mathbf{z}}\, \ell(d\mathbf{z})\,,$$

while the covariance matrix C remains the same.

Proof. Using Bayes' Theorem along with the density process of Q_θ and the independent increment property of the Lévy process, yield that the conditional log-characteristic function of $L(t)$ given $\mathcal{F}_s$ for $t \geq s \geq 0$ is

$$\ln \mathbb{E}_{Q_s}\left[\mathrm{e}^{\mathrm{i}\mathbf{x}'L(t)} \mid \mathcal{F}_s\right] = (\psi(\mathbf{x} - \mathrm{i}\theta) - \psi(-\mathrm{i}\theta))\,(t-s)\,.$$

Hence, L is a Lévy process under Q_θ as well. By a direct computation, we find the drift and the Lévy measure as claimed. □

Note that if we have a (bivariate) drifted Brownian motion as Lévy process, that is, $\ell(d\mathbf{z}) = 0$, then the Esscher transform is simply a Girsanov transform of the Brownian motion with a constant parameter θ. For Lévy processes with jumps, the Lévy measure is exponentially tilted by the Esscher transform. We may interpret this as a rescaling of the size and intensity of jumps.

We remark that the expected value of $L(1)$ under Q_θ is given by

$$\mathbb{E}_\theta\left[L(1)\right] = -\mathrm{i}\nabla\psi(-\mathrm{i}\theta)\,,$$

where ∇ is the gradient and $\mathbb{E}_\theta[\cdot]$ is the expectation operator with respect to the probability Q_θ. Thus, the Lévy process $\widetilde{L}(t) = L(t) + \mathrm{i}\nabla\psi(-\mathrm{i}\theta)t$ becomes a martingale under Q_θ as it has expectation zero. This means in particular that under Q_θ, the dynamics of X and Y are, respectively,

$$dX(t) = -\mathrm{i}\psi_x(-\mathrm{i}\theta)\,dt + d\widetilde{L}_1(t) \tag{5.3.3}$$

and

$$dY(t) = \{-\mathrm{i}\psi_y(-\mathrm{i}\theta) - \eta Y(t)\}\ dt + d\widetilde{L}_2(t)\,.$$

Here, we have used the notation ψ_x and ψ_y as the partial derivatives of ψ with respect to the two variables x and y, respectively. The solution $Y(s)$ at time $s \geq t$, conditioned on $Y(t)$, of this Ornstein-Uhlenbeck process is

$$Y(s) = Y(t)\mathrm{e}^{-\eta(s-t)} + \frac{-\mathrm{i}\psi_y(-\mathrm{i}\theta)}{\eta}(1 - e^{-\eta(s-t)}) + \int_t^s \mathrm{e}^{-\eta(u-t)}\,\widetilde{L}_2(du) \tag{5.3.4}$$

Next, we consider pricing of swaps in this market. Let us start with analysing the implied swap price dynamics for the arithmetic model. The following result holds:

Proposition 5.3.2. *The swap price $F(t, T_1, T_2)$ is given by*

$$\begin{aligned} F(t, T_1, T_2) = \bar{\Lambda}(T_1, T_2) + X(t) + Y(t)\bar{\eta}(t, T_1, T_2) \\ - \frac{1}{2}\mathrm{i}\psi_x(-\mathrm{i}\theta)(T_2 - T_1) - \mathrm{i}\psi_x(-\mathrm{i}\theta)(T_1 - t) \\ + \frac{-\mathrm{i}\psi_y(-\mathrm{i}\theta)}{\eta}(1 - \bar{\eta}(t, T_1, T_2))\,, \end{aligned}$$

where

$$\bar{\eta}(t, T_1, T_2) = \frac{1}{T_2 - T_1}\int_{T_1}^{T_2} \mathrm{e}^{-\eta(s-t)}\,ds$$

and $\bar{\Lambda}(T_1, T_2)$ is the average value of the seasonality function $\Lambda(s)$ over the interval $[T_1, T_2]$.

Proof. From the expression in (5.3.3), we find (for $s \geq t$)

$$\mathbb{E}_{Q_\theta}[X(s)|\mathcal{F}_t] = X(t) - \mathrm{i}\psi_x(-\mathrm{i}\theta)(s - t)\,,$$

after appealing to the independent increment property of the Q_θ-Lévy process $\widetilde{L}_1$ with zero mean, and the $\mathcal{F}_t$-measurability of $X(t)$. Similarly, from the independent increment property of the Q_θ-Lévy process $\widetilde{L}_2$, having mean zero, we find from (5.3.4)

$$\mathbb{E}_{Q_\theta}[Y(s)|\mathcal{F}_t] = Y(t)\mathrm{e}^{-\eta(s-t)} - \frac{\mathrm{i}\psi_y(-\mathrm{i}\theta)}{\eta}(1-\mathrm{e}^{-\eta(s-t)}) .$$

Since

$$F(t,T_1,T_2) = \frac{1}{T_2-T_1}\int_{T_1}^{T_2} \{\Lambda(s) + \mathbb{E}_{Q_\theta}[X(s)+Y(s)|\mathcal{F}_t]\}\, ds$$

the result follows after using the Fubini Theorem. □

We note that $\bar{\eta}$ is the average value of the "volatility function" $\exp(-\eta(s-t))$ over the delivery period $[T_1,T_2]$, and takes the form

$$\bar{\eta}(t,T_1,T_2) = \frac{1}{\eta(T_2-T_1)}\left(\mathrm{e}^{-\eta(T_1-t)} - \mathrm{e}^{-\eta(T_2-t)}\right) ,$$

or,

$$\bar{\eta}(t,T_1,T_2) = \mathrm{e}^{-\eta(T_1-t)}\frac{1}{\eta(T_2-T_1)}\left(1-\mathrm{e}^{-\eta(T_2-T_1)}\right) . \tag{5.3.5}$$

In the representation (5.3.5), T_1-t is time left until start of delivery, and T_2-T_1 is length of delivery. We recognize the exponential damping factor $\exp(-\eta(T_1-t))$ as the Samuelson effect on the volatility, that is, the volatility of the spot is increasing as time to start of delivery is decreasing. The classical Samuelson effect says that the volatility of the forward price is exponentially increasing in time to maturity to the spot volatility (see Samuelson [52] and Benth *et al.* [19]). We note here that $\bar{\eta}(t,T_1,T_2)$ is not converging to the "spot volatility", being one in this context, but to a value less than this. The delivery period creates this violation of the classical Samuelson effect. It is natural from a financial and empirical point of view that the volatility of the electricity forward price is not converging to that of the spot as the forward price is the average of the spot over a delivery period.

We derive the dynamics of F in the next proposition

Proposition 5.3.3. *The Q_θ dynamics of the swap price is*

$$dF(t,T_1,T_2) = d\widetilde{L}_1(t) + \overline{\eta}(t,T_1,T_2)\, d\widetilde{L}_2(t) .$$

Proof. Since $\overline{\eta}'(t,T_1,T_2) = \eta\overline{\eta}(t,T_1,T_2)$, the result follows after applying the Itô formula for jump processes and the Q_θ-dynamics of X and Y. □

As is apparent from the definition of $F(t,T_1,T_2)$, it is a Q_θ-martingale process for $t \leq T_1$. Thus, it defines an arbitrage-free model for the stochastic evolution of electricity forward prices.

5.4 An empirical study of EEX spot and forward prices

In this Section we want to estimate the parameters in the spot model, and calibrate it to forward prices where we derive the market price of risk θ. It turns out that a joint estimation of spot and forward is most efficient, where one can make use of the asymptotic behaviour of forward prices to filter out the non-stationary factor in the spot. This approach is analogous of the calibration procedure in Schwartz and Smith [55], with a more sophisticated version of it found in Benth *et al.* [16].

The following asymptotic result of the forward price with respect to time to delivery plays a crucial role in the estimation algorithm.

Proposition 5.4.1. *It holds that*

$$\lim_{T_1 - t \to \infty} \left\{ F(t, T_1, T_2) - \overline{\Lambda}(T_1, T_2) - \Psi(t, T_1, T_2; \theta) - X(t) \right\} = 0\,,$$

where

$$\Psi(t, T_1, T_2; \theta) = -\frac{1}{2}\mathrm{i}\psi_x(-\mathrm{i}\theta)(T_2 - T_1) - \mathrm{i}\psi_x(-\mathrm{i}\theta)(T_1 - t) - \frac{\mathrm{i}\psi_y(-\mathrm{i}\theta)}{\eta}\,.$$

Proof. Recalling the explicit dynamics of $F(t, T_1, T_2)$ in Proposition 5.3.2, the result follows after observing that $\exp(-\eta(T_1 - t)) \to 0$ as $T_1 - t \to \infty$. □

Hence, asymptotically the forward price behaves like

$$F(t, T_1, T_2) \approx \overline{\Lambda}(T_1, T_2) + \Psi(t, T_1, T_2; \theta) + X(t)\,, \tag{5.4.1}$$

for $T_1 - t \to \infty$. This means that in the long end of the forward market, the prices fluctuate as the non-stationary factor $X(t)$ plus some non-stochastic adjustment term $\bar{\Lambda}(T_1, T_2) + \Psi(t, T_1, T_2; \theta)$ involving the market price of risk θ. From these considerations we can derive an algorithm for estimating the model. It goes as follows.

For a fixed delivery period $[T_1, T_2]$,

1. Fit a seasonal function $\Lambda(t)$ to the spot prices $S(t)$.

2. Fit the autocorrelation function of $Y(t)$ to the de-seasonalized spot prices to have an *apriori* estimate of η. Use this η to find a threshold $\widehat{T}$ for which "$T_1 - t = \infty$", that is, how big should $T_1 - t$ be for the asymptotic behaviour of F in (5.4.1) to be acceptable.

3. Subtract $\overline{\Lambda}(T_1, T_2)$ from the observed forward prices to "de-seasonalize" them. Call this time series $\widetilde{F}(t, T_1, T_2)$

4. Observe that we have for $T_1 - t \geq \widehat{T}$

$$\widetilde{F}(t, T_1, T_2) \approx c(\theta, T_1, T_2) - \mathrm{i}\psi_x(-\mathrm{i}\theta)(T_1 - t) + X(t)$$

where

$$c(\theta, T_1, T_2) = -\frac{1}{2}\mathrm{i}\psi_x(-\mathrm{i}\theta)(T_2 - T_1) - \frac{\mathrm{i}\psi_y(-\mathrm{i}\theta)}{\eta}$$

Hence, for all observed forward prices $F(t, T_1, T_2)$ for which $T_1 - t \geq \widehat{T}$, estimate the "constants" $c(\theta, T_1, T_2)$ and $-\mathrm{i}\psi_x(-\mathrm{i}\theta)$ by linear regression of $\widetilde{F}$ with respect to $T_1 - t$.

5. Using the estimated regression coefficients $\widehat{c}$ and $\widehat{a}$, we filter out $X(t)$ from the observations,

$$\widetilde{F}(t, T_1, T_2) - \widehat{c} - \widehat{a}(T_1 - t)$$

for all $T_1 - t \geq \widehat{T}$.

6. Subtract the filtered data series $X(t)$ from the deseasonalized spot prices. This results in a time series which is modelled by $Y(t)$. Re-estimate η based on linear regression of $Y(t)$ against $Y(t-1)$.

7. Fit a Lévy process L to the residuals of the Y process and the time series of the X process obtained above. From the fitted Lévy process L, we obtain the cumulant ψ.

8. For the given cumulant ψ, find the estimated market price of risk θ by solving the system of equations

$$\begin{aligned}\widehat{a} &= -\mathrm{i}\psi_x(-\mathrm{i}\theta) \\ \widehat{c} &= -\frac{1}{2}\mathrm{i}\psi_x(-\mathrm{i}\theta)(T_2 - T_1) + \frac{-\mathrm{i}\psi_y(-\mathrm{i}\theta)}{\eta}\end{aligned} \tag{5.4.2}$$

This calibration algorithm provides us with a full specification of both the spot and the forward price model, including the estimation of the market prices of risk $\theta = (\theta_1, \theta_2)$. We next apply it to spot and forward price data collected from the European Energy Exchange (EEX).

We had available daily Phelix base load spot prices from 02.01.2006 until 19.10.2008, constituting altogether 1022 daily observations. Remark that we include weekend prices as we are going to apply base load forward prices in our estimation routine. These futures are settled on the spot prices including the weekends. To the spot price data, we fit the seasonality function taken from Barndorff-Nielsen *et al.* [7],

$$\Lambda(t) = \xi_0 + \xi_1 \cos(\frac{\tau_1 + 2\pi t}{365}) + \xi_2 \cos(\frac{\tau_2 + 2\pi t}{7}) + \xi_3 t + \xi_4 \mathbf{1}_{\mathrm{Sat}}(t) + \xi_5 \mathbf{1}_{\mathrm{Sun}}(t)\,.$$

This function takes annual and weekly seasonality into account along with a trend. As prices on weekends are in general lower than during the rest of the week due to a different demand situation, we introduce additionally a weekend-correction to capture these effects. Here $\mathbf{1}_{\mathrm{Sat}}(t)$ and $\mathbf{1}_{\mathrm{Sun}}(t)$ are equal 1, if the weekday corresponding to t is a Saturday and Sunday, respectively.

A non-linear least squares estimation on the spot data yields the parameters reported in Table 5.1.

Fig. 5.1 (left) displays the spot price data and its estimated seasonality function. The estimated seasonality follows the general movements of the spot, on a weakly pattern as well as a yearly one.

Next we continue the calibration algorithm with filtering the non-stationary factor X from the forward data with long time to delivery. For this purpose we use base load forward

ξ_0	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	τ_1	τ_2
738.733	4.360	-11.716	0.020	1.000	1.000	-13637.760	40.401

Table 5.1: Estimated parameters of the seasonal function.

contracts with 1 month delivery period from the EEX, for which we had available price data for the same dates as the spot (weekends and holidays are excluded, as there is no trade in futures).

We first need to determine the threshold $\widehat{T}$ for which the forward prices are asymptotically given by (5.4.1). This depends, obviously, on the value of η, the speed of mean reversion in the factor process Y. We can estimate this parameter from the autocorrelation function of Y which is known to be exponentially decaying at the rate η (see Benth *et al.* [19]). However, at this point in the estimation procedure we have not yet filtered the time series of Y from the spot data, so the empirical autocorrelation function is unknown to us. On the other hand, we do a rough estimation of η by looking at the empirical autocorrelation of the deseasonalized spot, which is modelled by $X(t) + Y(t)$. We observed a decaying autocorrelation structure, and fitted an exponentially decaying function to the first five lags obtaining the pre-estimate $\widehat{\eta} = 0.1781$. We derived $\widehat{T} = 16$ as the threshold when $Y(t)\overline{\eta}(t, T_1, T_2) \approx 1$ using $Y(t)$ being three times the standard deviation of spot price data. Note that we expect the presence of X to make the beta smaller than the "true" one. A larger value for η would lead to a smaller threshold. Hence, our decision to apply $\widehat{T} = 16$ is a conservative choice.

We construct a time series of forward prices with "infinite" time to delivery from the base load contracts as follows: if the time to delivery is more than 16 days, we choose at time t the forward with delivery period the first coming month. Otherwise, we switch to the contract with delivery in the following month. That is, we use the price series of front-month contracts as long as these are farther than 16 days to delivery, and switch to the next month when the front-month contracts have less than 16 days to delivery. Like this we make sure that for each date we have a forward price with time to delivery more than 16 days. These prices will not, at least approximately, have any influence from the stationary component Y. As the futures are not traded on weekends and holidays, we use as a substitute for missing values in the weekend the price on the preceding Friday. In holidays, we use the price on the last trading day before the holidays.

To deseasonalize the constructed forward price series we subtract the average seasonality of the delivery period. We have fitted the seasonality function to data until October 2008, such that we take October 2008 as the last delivery period and let our forward price series end at 14.09.2008. A linear regression of this time series delivers the estimates $\hat{a} = 0.030$ and $\hat{c} = 3.406$. We filter the non-stationary time series $X(t)$ from the forward prices corresponding to step (5) in the algorithm, and afterwards retrieve the stationary time series $Y(t)$ from the spot prices as in step (6). The plot on the right in Fig. 5.1 shows the filtered factor $X(t)$ along with the deseasonalized spot prices. It seems to reflect a long-term stochastic trend in the price data.

Next we estimate the mean reversion parameter η. The autocorrelation function of the time series $Y(t)$ is plotted in Fig. 5.2. Re-estimating η over the first five lags results in

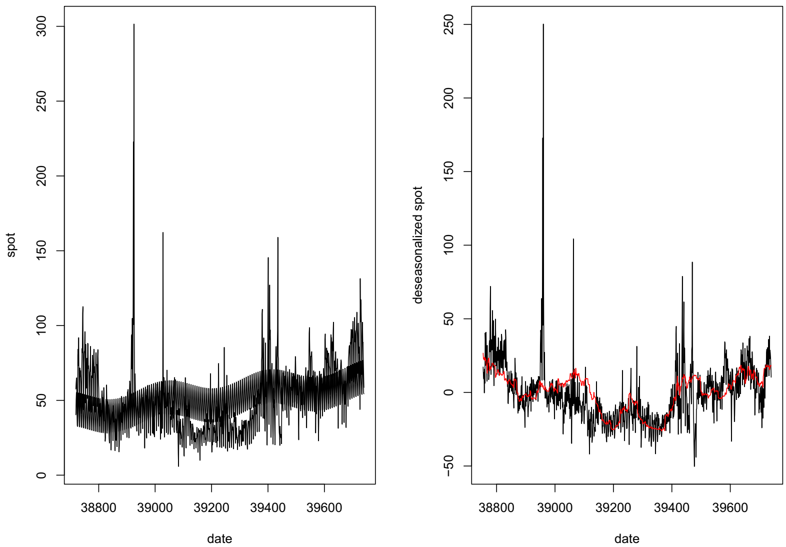

Figure 5.1: Left: empirical spot price data together with the estimated seasonality function. Right: deseasonalized spot price data with the filtered data series $X(t)$

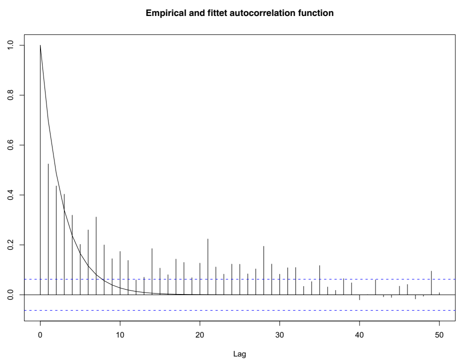

Figure 5.2: Autocorrelation function of $Y(t)$

$\widehat{\eta} = 0.359$. The initial decrease of the autocorrelation function seems to be captured well by using an exponential function. However, it decays too rapidly for larger lags. Including more lags to fit the autocorrelation function (that is, η) results in a poor fit in the first lags. To get a better fit over all lags, one could use two (or more) exponential components. This would mean that we model the factor Y by two or more Ornstein-Uhlenbeck processes, or by a higher-order CARMA model. Benth *et al.* [16] indicate that one should indeed use a higher-order CARMA model. However, such models are much more complex to estimate,

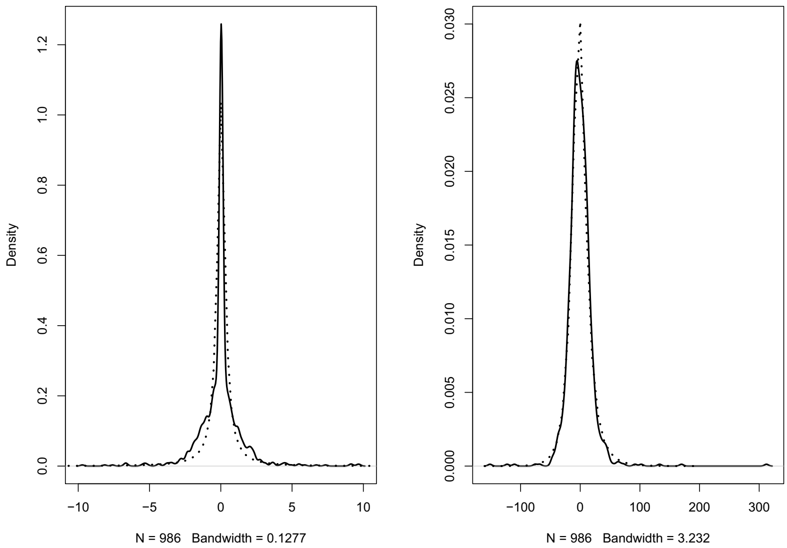

Figure 5.3: empirical density of L_1 (left) and L_2 (right) as well as the fitted NIG density (dashed line)

and we apply the one-factor assumption on Y here as a first approximation of the dynamics.

The next step is to fit a bivariate Lévy process $L = (L_1, L_2)$ to the time series $X(t)$ and $Y(t)$. For simplicity, we assume that L_1 and L_2 are independent, meaning that there is no dependency between the short-term and long-term price fluctuations. In the Schwartz-Smith model (see Schwartz and Smith [55], or Lucia and Schwartz [47] for the case of electricity) L is assumed to be a bivariate Brownian motion. However, the Gaussian assumption on the increments $\Delta X(t)$ is not realistic, and we propose to fit the dynamics of X with a normal inverse Gaussian (NIG) Lévy process, that is, a Lévy process with NIG distributed marginals. The NIG distribution seems to be a good choice for modelling the residuals of $Y(t)$ as well.

The NIG distribution is a four parameter family of distributions successfully applied to model the logreturns of financial data. For its applications to finance and a detailed probabilistic analysis of the NIG family, we refer the interested reader to Barndorff-Nielsen [6]. Assuming $L_1(t)$ to be a NIG Lévy process, its cumulant (*i.e.*, the logarithm of the characteristic function) function at time 1 is given by

$$\Psi(x) = \delta\big\{\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + ix)^2}\big\} + \mu ix\,, \tag{5.4.3}$$

for the four parameters μ, β, $\delta > 0$ and $\alpha > 0$. The skewness of the NIG distribution is described by β, where $\beta > 0$ means a positively skewed distribution, and $\beta < 0$ negatively skewed. For a symmetric NIG distribution, *i.e.*, when $\beta = 0$, μ is the mean. Otherwise, μ is the location parameter. δ is the scale and α the tail heaviness parameter. Note that the NIG distribution has semi-heavy tails, with the normal distribution as a limiting case. We easily

	α	β	δ	μ
L_1	0.0946	−0.0099	0.3136	0.02421
L_2	0.0402	0.0071	14.3407	−2.9488

Table 5.2: Estimated NIG parameters of L_1 and L_2.

find the expectation from (5.4.3) as

$$\kappa_1 = \frac{\delta\beta}{\sqrt{\alpha^2 - \beta^2}} + \mu\,.$$

The estimated parameters of $L_1(1)$ based on maximum-likelihood are given in Table 5.2. We remark in passing that the NIG distribution has been applied in studies of energy prices in Benth and Šaltytė-Benth [18] and Börger *et al.* [23].

We fit another NIG Lévy process L_2 to the residuals of Y. The estimates are reported in Table 5.2. The estimated densities of $L_1(1)$ and $L_2(1)$ are displayed together with the empirical ones in Fig. 5.3. The fit seems to be good, and we find the NIG distribution as a satisfactory choice for modelling L_1 and L_2. Recall that we assumed independence of L_1 and L_2. Empirically, the correlation between the data series for L_1 and L_2 is given by -0.16. A more realistic model should take this into account, which requires an analysis of the dependency structure. We relegate this to future studies. From the estimates in Table 5.2 we observe that the NIG distributions for L_1 and L_2 are close to symmetric.

Following step (8), the market price of risk $\theta = (\theta_1, \theta_2)$ is given by

$$\begin{aligned}\theta_1 &= \frac{\alpha_1 \frac{\hat{a}-\mu}{\delta}}{\sqrt{(\frac{\hat{a}-\mu_1}{\delta_1})^2 + 1}} - \beta_1 \\ \theta_2 &= \frac{\alpha_2 K}{\sqrt{K^2+1}} - \beta_2\,,\end{aligned} \tag{5.4.4}$$

where

$$K = \frac{\beta_2}{\delta_2}\left(\hat{c} - \frac{1}{2}\hat{a}(T_2 - T_1) - \frac{\mu_2}{\beta_2}\right)\,.$$

Here, the subscript in the parameters α, β, δ and μ refer back to L_1 and L_2. Using the estimates for the NIG distributions, we can derive the values of θ_1 and θ_2. These are reported in Table 5.3 along with the expected values of L_1 and L_2 with respect to the probabilities P and the fitted Q_θ. We note that the market price of risk is positive, and that the expected value of L_1 and L_2 are moved from being negative under P to positive under Q_θ. The fitted market price of risk is shifting the distribution of L_1 and L_2 towards the right, roughly meaning that we get more positive jumps and less negative. Furthermore, quite nicely the NIG distribution is preserved under a constant Esscher transform. Hence, L is a bivariate NIG Lévy process both under P and Q_θ, where only the skewness parameter is different under the two measures.

i	θ_i	$\mathbb{E}_P[L_i(1)]$	$\mathbb{E}_\theta[L_i(1)]$
1	0.0115	-0.0087	0.0296
2	0.0010	-0.3583	0.0211

Table 5.3: The market price of risk derived from the fitted NIG parameters together with the expectation of L_1 and L_2 under P and Q_θ

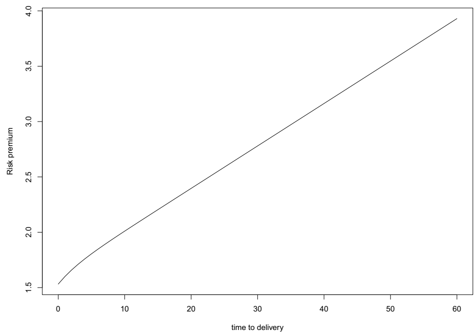

Figure 5.4: Theoretical risk premium for the estimated model parameters.

Let us comment on the risk premium implied by our estimated model. The risk premium is defined as the difference between the forward price and the predicted average spot price over the delivery period. In mathematical terms,

$$R_P(t, T_1, T_2) = F(t, T_1, T_2) - \mathbb{E}\left[\frac{1}{T_2 - T_1}\int_{T_1}^{T_2} S(t)\,dt \,|\, \mathcal{F}_t\right].$$

From Proposition 5.3.2 we find

$$\begin{aligned} R_(t, T_1, T_2) = {} & \frac{1}{2}(\mathbb{E}_\theta[L_1(1)] - \mathbb{E}[L_1(1)])(T_2 - T_1) \\ & + (\mathbb{E}_\theta[L_1(1)] - \mathbb{E}[L_1(1)])(T_1 - t) \\ & + (\mathbb{E}_\theta[L_2(1)] - \mathbb{E}[L_2(1)])\frac{1}{\eta}(1 - \bar{\eta}(t, T_1, T_2)). \end{aligned}$$

The non-stationary factor gives a linear contribution in time to delivery $T_1 - t$, while the stationary factor gives an exponential shape and converges fastly to a constant when $T_1 - t \to \infty$. A plot of the risk premium for the estimated model parameters is shown in Fig. 5.4. As a result of the positive market price of risk, the risk premium also becomes positive. This tells that the consumers in the market are willing to pay a premium for locking in electricity prices in the forward market. Note that we use data from the relative short end of the market, using the front-month (or second month) contracts.

Contract	Trading day	Delivery period	Strike	Futures price	Settlement price
C1	06.02.2008	Mar 2008	57	56.81	1.900
C2	28.01.2008	Mar 2008	57	57.00	2.270
C3	15.01.2008	Feb 2008	75	70.50	1.065
C4	09.01.2008	Feb 2008	74	68.50	0.928

Table 5.4: Traded call options in 2008 with delivery period 1 month

Contract	Trading day	Delivery period	Strike	Futures price	Settlement price
P1	08.07.2008	Aug 2008	74	74.77	3.233
P2	08.07.2008	Aug 2008	75	74.77	3.835
P3	03.07.2008	Aug 2008	73	78.00	1.989
P4	08.04.2008	May 2008	55	55.35	1.522
P5	04.03.2008	Apr 2008	58	58.70	1.911
P6	28.02.2008	Apr 2008	58	61.75	0.955
P7	08.01.2008	Feb 2008	65	69.00	1.179

Table 5.5: Traded put options in 2008 with delivery period 1 month

5.5 Pricing of options on futures

At EEX, the market for options is rather illiquid, however, there exists traded contracts. In 2008, 12 options on baseload futures with delivery period 1 month were traded, 11 of them in the period we consider. Out of these 11, four are call options, and seven puts. We use these for further analysis and discussion.

In Tables 5.4 and 5.5 we list the calls and puts with their main characteristics. We have decided to label the contracts by Ci, $i = 1, 2, 3, 4$ for the calls and Pi, $i = 1, \ldots, 7$ for the puts. Recall that the exercise time τ of the options is four trading days before the delivery period of the underlying futures starts. The historical data available from the EEX provides settlement prices for traded option contracts. For all derivatives traded, a settlement price is established on all exchange trading days. In the case that a settlement price cannot be determined on basis of the order book situation, a so-called Chief Trader Procedure applies, where all trading participants can take part with a representative. The EEX Market Supervision makes a standardised form available for all those trading participant volunteering to specify a market price for the respective derivatives. The settlement price is then determined as the average of the expectations of the market participants. We note that options on peak-load futures are not traded at all at EEX, explaining why we use baseload spot data in our empirical analysis above.

We first look at the "classical" approach to pricing options on futures in commodity markets, namely pricing using the Black-76 formula. For the convenience of the reader, we state the Black-76 formula in a proposition.

Contract	Settlement price	Black-76	Mispricing	hist. vol.	impl. vol.
C1	1.900	0.464	-76%	0.1046	0.3770
C2	2.270	0.725	-68%	0.1100	0.3560
C3	1.065	0.000	-100%	0.0788	0.5030
C4	0.928	0.000	-100%	0.0821	0.4450

Table 5.6: Black-76 pricing of the call options

Proposition 5.5.1. *Suppose the risk-neutral futures price dynamics is a geometric Brownian motion*

$$\frac{dF(t,T_1,T_2)}{F(t,T_1,T_2)} = \sigma\, dB(t)\,,$$

for a constant $\sigma > 0$. Then, the price at time $t \leq \tau$ of a call option with strike K and exercise time $t \leq \tau \leq T_1$, is given by $C_{B76}(t, F(t,T_1,T_2))$ with

$$C_{B76}(t,x) = e^{-r(\tau-t)}[x\Phi(d_1(x)) - K\Phi(d_2(x))]\,,$$

for Φ being the cumulative standard normal distribution function, and

$$d_1(x) = \frac{\ln\left(\frac{x}{K}\right) + \frac{1}{2}\sigma^2(\tau-t)}{\sigma\sqrt{\tau-t}}\,,$$
$$d_2(x) = d_1 - \sigma\sqrt{\tau-t}\,.$$

In the Black-76 formula, one boldly assumes the futures price dynamics to be geometric Brownian motion, a dynamics which is far from the one we have estimated to the electricity futures prices at the EEX. The volatility σ is also constant, an assumption that is not likely to be true. Based on the historically estimated volatility of the futures contracts in question, we can price the call options. The Black-76 prices are reported in Table 5.6 along with the actual settlement prices as quoted on the EEX. Appealing to the put-call parity, we report the put prices in Table 5.7. In both tables, we have also reported the historical volatility σ used in the Black-76 formula, as well as the implied volatility so that Black-76 matches the settlement price. We estimate the historical volatility of the logreturns of the underlying futures from the last month of daily price data. Furthermore, we choose $r = 5\%$ which is about the average yearly Euro LIBOR rate in 2008. We find that the price of all options are substantially underestimated by Black-76. Due to the low volatility, those options that are far out of the money have a Black-76 price being essentially 0 (P6 and P7, and C3 and C4). The implied volatility becomes very high compared to the historical volatility. Indeed, the historical volatility is in the modest range of 8-11% for the underlying futures of the call, whereas the implied volatilities are estimated to be from 35% to 50%. The mispricing is rather dramatic, as the percentages ranging above 70% tells. One might be tempted to speculate that the market is adding a huge risk premium for effects like illiquidity of the options and non-normality of the futures price dynamics. The issuer runs a big risk selling call options, since it is difficult to turn around the position in the option market. However,

Contract	Settlement price	Black-76	Mispricing	hist. vol.	impl. vol
P1	3.233	0.693	-79%	0.1491	0.521
P2	3.835	1.158	-70%	0.1491	0.532
P3	1.989	0.055	-97%	0.1496	0.509
P4	1.522	0.177	-88%	0.0679	0.357
P5	1.911	0.295	-85%	0.1014	0.394
P6	0.955	0.001	-100%	0.0797	0.366
P7	1.179	0.000	-100%	0.0842	0.437

Table 5.7: Black-76 pricing of the put options

the underlying futures is reasonably liquid, so delta hedging is possible. This removes some of the liquidity risk for the issuer.

One can in theory create synthetic investment strategies mimicking to a large extent the payoff of a call or put option. This could be used in order to exploit potential arbitrages in the option market. However, if the futures dynamics is not a geometric Brownian motion, there will be a large residual error in such strategies, which theoretically can be made perfect by delta hedging in the Black-76 framework. The empirical study of spot and futures pricing in the previous Section strongly points towards non-Gaussian models, hence ruling out this possibility.

In any case, the conclusion so far is that Black-76 in its simplest form is inadequate for pricing of options in the EEX market. As our proposed futures price dynamics is far more sophisticated than a simple geometric Brownian motion, we now move on to analyse the implied option prices from this model with the hope that it can improve the situation.

The call option price is then given by

$$C(t) = e^{-r(\tau-t)}\mathbb{E}_Q[F(\tau, T_1, T_2) - K \mid \mathcal{F}_t]\,. \tag{5.5.1}$$

The pricing probability Q is an equivalent martingale measure for $F(t, T_1, T_2)$, and we let this be given by Q_θ. The Q_θ-dynamics of $F(t, T_1, T_2)$ is given by Proposition 5.3.3 and Q_θ is determined through the market price of risk (5.4.4) from the spot-futures analysis above. We evaluate the expectation through a Monte-Carlo simulation. To simulate the Lévy processes L_1 and L_2 under Q_θ, we use that NIG-Lévy processes are stable with respect to an Esscher change of measure. In fact, it can be seen (see Benth *et al.* [19]) that if, for $i = 1, 2$, $L_i(1)$ is NIG distributed under P with parameters $\alpha_i, \beta_i, \delta_i$ and μ_i, then the $L_i(1)$ is again NIG distributed under Q_θ with the same parameters except the skewness, which becomes $\beta_i + \theta_i$.

Based on a simulation of 1,000,000 paths we compute the option prices based on the average payoff. To simulate the NIG distribution, we applied the algorithm implemented in the R-package fBasics, which is based on the normal variance-mean mixture of the NIG distribution.

The resulting numbers are reported in Tables 5.8 and 5.9. We have also included the misspricing and computed the implied volatility of the simulated price using the Black-76. From the tables, we see that the picture is more mixed, with both over and underpricing of the calls and puts. Moreover, at the first glance, the mispricing seems to be less severe than in the case of Black-76, although admittedly still very big.

Contract	Settlement price	Simulated price	Mispricing	impl. vol.
C1	1.900	2.748	45%	0.4820
C2	2.270	3.525	56%	0.4882
C3	1.065	0.821	-23%	0.4255
C4	0.928	1.006	9%	0.4037

Table 5.8: Simulated prices of the call options

Contract	Settlement price	Simulated price	Mispricing	impl. vol.
P1	3.233	2.476	-23%	0.4256
P2	3.835	2.964	-23%	0.4239
P3	1.989	1.438	-28%	0.4260
P4	1.522	2.397	57%	0.5358
P5	1.911	2.659	39%	0.5368
P6	0.955	1.889	98%	0.5384
P7	1.179	1.376	17%	0.4032

Table 5.9: Simulated prices of the traded put options

Our spot and futures price model includes non-Gaussian noise as both factors in the spot are driven by an NIG Lévy process. Note that the futures price is depending on the non-stationary factor directly, whereas the short-term factor is dampened and negligible for contracts far from delivery. From our estimation procedure, the non-stationary long-term factor is estimated from the futures prices, so if the market would price according to our futures price dynamics with the given pricing measure Q_θ, at least options with long time until exercise should be priced reasonably accurate. Looking at C1 and C2, these have the longest time to exercise in our sample of call options. However, the simulated option prices from our model for these two contracts are approximately 50% higher than the quoted prices. This means that our model is pricing in *too much* risk. From the spot dynamics we estimated positive market prices of risk which pushes the skewness of the NIG distribution to more positive jumps. The more positive market price of risk, the higher values of the options. Thus, it seems like the forward model inherits far too much risk premium from the spot when it comes to option pricing. We reach the conclusion that the option market is not including the same risk perception as the one inherited from the spot in the futures market. This is a clear sign that a completely different pricing measure Q is used in the option market than in the futures pricing. Note that C1 and C2 are both (approximately) at-the-money, so a big portion of the distribution of the futures is taken into account in the pricing.

The contracts C3 and C4 are far out-of-the-money and slightly closer to exercise time than C1 and C2. Noteworthy is that the mispricing of these are significantly less than for C1 and C2, being respectively -23% and 9%. If we have based our calculations of the call prices on the wrong risk premium, it will be more influential far from exercise than close since we span out more of the risk the longer into the future we simulate the futures price. Close to exercise, the mis-specification of the tails under the chosen Q will be relatively much smaller than when we move forward. Maybe more importantly is that a smaller portion of the price distribution of F is taken into account for these two out-of-the-money options than C1 and C2, and hence a wrongly chosen Q matters less. This discussion conforms with the observations above for C1 and C2.

Note that contract P6 is farthest from exercise among the put options, as well as being out-of-the-money. This contract has the highest mis-pricing by our model. All the other put contracts in our sample have shorter time left to exercise. P1, P2, P4 and P5 are all approximately at-the-money put options with almost the same time left until exercise. The mis-pricing of these are significantly less than for P6. In fact, for P1 and P2 our model gives a price -23% less than the settlement prices. P4 and P5 are contracts on futures with delivery in the spring months May and April, respectively. Temperature predictions may influence the futures price expectations, as the spring may become colder or warmer than usual. We also note that it is the left-tail of the futures price distribution that counts when pricing an out-of-the-money out option. An under-pricing can be the result of the distribution being moved to the right by a risk positive premium.

P3 and P7 are out-of-the-money put options where the mis-pricing of our model is rather modest (-28% and 17%, respectively). P7 is the only put option written on a futures with delivery in the winter period, namely February. For the calls C3 and C4, which also are written on February futures contacts, we observe a relatively small pricing error. It seems like the model captures best the futures price evolution in the winter term. We also remark that the poor fit of the autocorrelation of the stationary factor Y may lead to wrong assessments of the spike influence. However, we believe that this is to some extent compensated for by the good fit of the Lévy process L_2 using a NIG distribution.

All in all, it seems like the futures price dynamics based on the pricing measure Q_θ implied by the spot-futures relationship provides a significantly better prediction of option prices than Black-76. However, the prices are far from satisfactory, and we find clear evidence that the risk-adjustments should be different than those given by Q_θ. Based on our findings, we dare to conclude that another pricing measure $\widetilde{Q}$ should be used for power option pricing, a pricing measure which attributes a different loading on the distributions of the Lévy processes L_1 and L_2. In fact, based on the differences between summer and winter contracts in the pricing analysis above, one may suspect that such a measure change should incorporate seasonalities as well. Furthermore, it may also account for the state of the futures price, so that one can capture out- or and in-the-money option price differences better. One can also think of pricing measures which not only changes the characteristics of the jump processes L_1 and L_2, but as well change the dynamics. For example, it is possible to define measures which change the speed of mean reversion of the Y factor. This could for example lead to a slower risk-neutral speed of mean reversion, essentially saying that a spike lasts longer in a risk neutral context than under the market probability.

As our futures price dynamics consists of two jump components, it gives rise to an incomplete market model. The selection of risk neutral probabilities for pricing options in such markets is frequently based on utility indifference pricing techniques (see Rouge and El Karoui [50]). Such a method, which is based on a risk averse, utility optimizing investor, leads to a partial hedging strategy of the option. The utility indifference method is particularly useful when pricing options in illiquid markets, where one is stuck with the option investment. Other approaches to pricing is based on deriving optimal partial hedges, where the optimality criterion may be the futures investment hedge which minimizes the variance of the hedging error (see Cont and Tankov [31]). All these various approaches lead to a pricing measure Q. It is of great interest and application to see whether such prices will explain the settlement option prices in the EEX market, and whether our conjecture $\widehat{Q} \neq \widetilde{Q}$ is true.

5.6 Conclusion

We have argued that in power markets one may use a probability measure $\widehat{Q}$ for futures pricing based on spot modelling which can be different than the equivalent martingale measure $\widetilde{Q}$ used for pricing options on the futures. There is no violation of no-arbitrage pricing theory that $\widehat{Q} \neq \widetilde{Q}$, and the argument hinges on the fact that electricity spot cannot be stored. Due to the non-storability, $\widetilde{Q}$ can be chosen as an equivalent measure which is not necessarily turning the discounted spot dynamics into a $\widetilde{Q}$-martingale. On the other hand, $\widehat{Q}$ is an equivalent measure such that the futures price becomes a $\widehat{Q}$-martingale.

We introduce a two-factor model for the spot price dynamics being a generalization of the classical commodity model of Schwartz and Smith [55]. Both the long-term and the short-term factors are driven by normal inverse Gaussian Lévy processes, a choice based on empirical arguments using data collected at the EEX. The spot model allows for analytical futures pricing, where the Esscher transform provides an parametric class of probability measures to model the risk premium. We perform a joint estimation of the spot and forward, where the crucial step is to apply long-dated futures contracts to filter out the non-stationary long-term factor of the spot.

Applying Monte Carlo simulations we priced call and put option prices for our proposed futures dynamics. We compared the simulated prices where we chose $\widehat{Q} = \widetilde{Q}$ with observed option prices in the market. This led to a significant mis-pricing, and we argued that the results pointed to the fact that $\widetilde{Q} \neq \widehat{Q}$. Our results were benchmarked against the Black-76 prices using the historical volatility of the underlying futures as input. The proposed spot-futures model was a clear improvement over this in predicting option prices.

We did not suggest any probability $\widehat{Q}$ for the futures price which could remedy the situation. There exists many potential approaches to produce such risk neutral probabilities taken from the theory of derivatives pricing in incomplete markets. But before setting off such a study, one should make the spot dynamics even more sophisticated to take into account some defiances like the mis-specification of the autocorrelation structure of the stationary factor. CARMA processes could be a choice here, or even more general Lévy semistationary processes. However, this will require more advanced estimation procedures to fit to data. On the other hand, such improvements will make the conclusions on option pricing and choice of risk neutral measures less prone to model error. A further issue is to open for more flexible pricing measures for the futures price, taking into account random changes in the risk premium and impacts from fuels and weather.

Illiquidity of the power option markets is a clear issue which can question our analysis. Power options are relatively little traded, but we believe that in the future these markets will emerge as important one for hedging and speculation of power. The results in our paper will hopefully provide an important guideline in the challenges when pricing spot, futures and options simultaneously.

BIBLIOGRAPHY

[1] Benth, F. E., and Schmeck, M. D. (2012). Stability of Merton's Portfolio Optimization Problem for Lévy Models. *To appear in: Stochastics*, DOI:10.1080/17442508.2012.665056.

[2] Benth, F. E., Di Nunno, G., Khedher, A., and Schmeck, M. D. (2012). On stability to model risk of options in a bivariate Lévy market. *Preprint.*

[3] Benth, F. E., and Schmeck, M. D. (2012). Pricing and Hedging Options in Energy Markets by Black-76. *Submitted.*

[4] Benth, F. E., and Schmeck, M. D. (2012). Pricing Futures and Options in Electricity Markets. *Submitted.*

---

[5] Asmussen, S., and Rosinski, J. (2001). Approximations of small jumps of Lévy processes with a view towards simulation. *J. Appl. Prob.* **38**, pp. 482–493.

[6] Barndorff-Nielsen, O. E. (1998). Processes of normal inverse Gaussian type. *Finance Stoch.*, **2**(1), pp. 41–68.

[7] Barndorff-Nielsen, O. E., Benth, F. E., Veraat, A. E. D. (2011). Modelling energy spot prices by Lévy semistationary processes. *Preprint*.

[8] Barndorff-Nielsen, O. E., Benth, F. E., and Veraart, A. E. D. (2011). Modelling electricity forward markets by ambit fields. *Submitted manuscript*.

[9] Benth, F. E., Di Nunno, G., and Khedher, A. (2011). Robustness of option prices and their deltas in markets modelled by jump-diffusions. *Comm. Stochastic Analysis*, **5**(2), pp. 285–307.

[10] Benth, F. E., Di Nunno, G., and Khedher, A. (2010). A note on convergence of option prices and their Greeks for Lévy models. *E-print series in pure mathematics*, University of Oslo, 18, 2010.

[11] Benth, F. E., Di Nunno, G., and Khedher, A. (2010). Lévy-models robustness and sensitivity. In: Ouerdiane, H., and Barhoumi A. (eds.), *Quantum Probability and White Noise Analysis*, World Scientific, pp. 153–184.

[12] Benth, F. E., Di Nunno, G., Khedher, A. and Schmeck, M. D. (2012). Pricing of Spread Options on a Bivariate Jump Market and Stability to Model Risk. *E-print series in pure mathematics*, University of Oslo, 2, 2012. ISSN 0806-2439.

[13] Benth, F. E., Karlsen, K. H., and Reikvam, K. (1999). Optimal portfolio selection with consumption and nonlinear integro-differential equations with gradient constraint: A viscosity solution approach. *Finance Stochast.* **5**, pp. 275–303.

[14] Benth, F. E., Karlsen, K. H., and Reikvam, K. (2001). Optimal portfolio management rules in a non-Gaussian market with durability and intertemporal substitution. *Finance Stochast.*, **5**, pp. 447–467.

[15] Benth, F. E., Kiesel, R., and Nazarova, A. (2012). A critical empirical study of three electricity spot price models. *Energy Economics*, **34**(5), pp. 1589–1616.

[16] Benth, F. E., Klüppelberg, C., Müller, G., and Vos, L. (2012). Futures pricing in electricity markets based on stable CARMA spot models. *Submitted manuscript*.

[17] Benth, F. E., and Koekebakker, S. (2008). Stochastic modeling of financial electricity contracts. *Energy Economics*, **30**(3), pp. 1116–1157.

[18] Benth, F. E., and Šaltytė-Benth, J. (2004). The normal inverse Gaussian distribution and spot price modelling in energy markets. *Intern. J. Theor. Appl. Finance*, **7**(2), pp. 177–192.

[19] Benth, F. E., Šaltytė-Benth, J., and Koekebakker, S. (2008). *Stochastic Modelling of Electricity and Related Markets*. World Scientific.

[20] Benth, F. E., and Sgarra, C. (2012). The risk premium and the Esscher transform in power markets. *Stoch. Analysis Appl.*, **30**, pp. 20–43.

[21] Black, F. (1976). The pricing of commodity contracts. *J. Financial Economics*, **3**, pp. 169–179.

[22] Black, F., and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*, **81**(3), pp. 637–654.

[23] Börger, R., Cartea, A., Kiesel, R., and Schindlmayer, G. (2009). A multivariate commodity analysis and applications to risk management. *J. Futures Markets*, **29**(3), pp. 197–217.

[24] Borovkova, S. A. (2008). Forward Curve Modelling in Commodity Markets. In: German, H. (eds.), *Risk management for commodities: from shipping to agriculture and energy*, The Wileys Finance Series, pp. 9-32.

[25] Borovkova, S. A., Permana, F. J., v.d. Weide, H. (2007). A closed form approach to the valuation and hedging of basket and spread options. *J. of Derivatives*, **14**(4) pp. 8–24.

[26] Bølviken, E., and Benth, F. E. (2000). Quantification of risk in Norwegian stocks via the normal inverse Gaussian distribution. Proceedings of the 10th AFIR colloquium, Tromsø 100, R. Norberg et al (eds.), pp. 87–98.

[27] Callegaro, G., and Vargiolu, T. (2009). Optimal portfolio for HARA utility functions in a pure jump multidimensional incomplete market. *Int. J. Risk Assessment and Management*, **11**, pp. 180–200.

[28] Carmona, R., and Durrleman, V. (2003). Pricing and hedging spread options. *Siam Review*, **45**(4), pp. 627–685.

[29] Cheang, G. H. L., and Chiarella, C. (2011). Exchange options under jump-diffusion dynamics. *Appl. Math. Finance*, **18**(3), pp. 245–276.

[30] Cohen, S., and Rosinski, J. (2005). Gaussian approximation of multivariate Lévy processes with applications to simulation of tempered and operator stable processes. *Bernoulli*, **13**, pp. 159–210.

[31] Cont, R., and Tankov, P. (2004). *Financial Modelling with Jump Processes*. Chapman & Hall/CRC.

[32] Delong, L., and Klüppelberg, C. (2008). Optimal investment and consumption in a Black-Scholes market with stochastic coefficients driven by a non-Gaussian Ornstein-Uhlenbeck process. *Ann. Appl. Probab.*, **18**(3), pp. 879–908.

[33] Duffie, D. (1992). *Dynamic Asset Pricing Theory*. Princeton University Press.

[34] Eberlein, E., Glau, K., and Papapantoleon, A. (2009). Analysis of Fourier transform valuation formulas and applications. *Appl. Math. Finance*, **17**(3), pp. 211–240.

[35] Eberlein, E., and Keller, U. (1995). Hyperbolic distributions in finance. *Bernoulli*, **1**, pp. 281–299.

[36] Eberlein, E., Papapapantoleon, A., and Shiryaev, A. N. (2009). Esscher transform and the duality principle for mulitdimensional semimartingales. *Ann. Appl. Prob.*, **19**(5), pp. 1944–1971.

[37] Emmer, S., and Klüppelberg, C. (2004). Optimal portfolios when stock prices follow an exponential Lévy process. *Finance Stochast.*, **8**, pp. 17–44.

[38] Folland, G. B. (1984). *Real Analysis – Modern Techniques and their Applications*. John Wiley & Sons.

[39] Garcia, I., Klüppelberg, C., and Müller, G. (2010). Estimation of stable CARMA models with an application to electricity spot prices. *Statist. Modelling*, **11**(5), pp. 447–470.

[40] Geman, H. (2005). *Commodities and Commodity Derivatives*. Wiley-Finance.

[41] Gerber, H. U., and Shiu, E. S. W. (1996). Actuarial bridges to dynamic hedging and option pricing. *Insurance: Math. Economics*, **18**, pp. 183–218.

[42] Gibson, R., and Schwartz, E. S., (1990). Stochastic convenience yield and the pricing of oil contingent claims. *J. Finance*, **45**, pp. 959–976.

[43] Jakobsen, E. R., Karlsen, K. H., and La Chioma, C. (2008). Error estimates for approximate solutions to Bellman equations associated with controlled jump-diffusions. *Numer. Math.*, **110**, pp. 221–255.

[44] Khedher, A. (2012). Computation of the Delta in Multidimensional Jump-Diffusion Setting with Applications to Stochastic Volatility Models. *Stochastic Analysis and Applications*, **30**(3), pp. 403–425.

[45] Klüppelberg, C., and Pergamenchtchikov, S. (2010). Optimal consumption and investment with bounded downside risk for power utility functions. In: Delbaen, F., Rásonyi, M., and Sticker, C. (eds.), *Optimality and Risk - Modern Trends in Mathematical Finance*, Springer, pp. 133–170.

[46] Larsen, K., and Žitković, G. (2007). Stability of utility-maximisation in incomplete markets. *Stoch. Proc. Appl.*, **117**(11), pp. 1642–1662.

[47] Lucia, J. J., and Schwartz E. S. (2002). Electricity prices and power derivatives: evidence from the Nordic Power Exchange. *Rev. Derivatives Res.*, **5**(1), pp. 5–50.

[48] Margrabe, W. (1978). The value of an option to exchange one asset for another. *J. Finance*, **33**, pp. 177–186.

[49] Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: the continuous-time case. *Rev. Econ. Statist.*, **51**, pp. 247–257.

[50] Rouge, R., and El Karoui, N. (2000). Pricing via utility maximization and entropy. *Math. Finance*, **10**, pp. 259–276.

[51] Rydberg, T. H. (1997). The normal inverse Gaussian Lévy process: simulation and approxoimation. *Commun. Statist. – Stochastic Models*, **13**, pp. 887–910.

[52] Samuelson, P. (1965). Proof that properly anticipated prices fluctuate randomly. *Indust. Manag. Rev.*, **6**, pp. 41–44.

[53] Sato, K. I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.

[54] Schilling, R. L. (2005). *Measures, Integrals and Martingales*. Cambridge University Press.

[55] Schwartz, E., and Smith, J. E. (2000). Short-term variations and long-term dynamics in commodity prices. *Manag. Science*, **46**(7), pp. 893–911.

[56] Shiryaev, A. N. (1999). *Essentials of Stochastic Finance*. World Scientific.

[57] Pasin, L., and Vargiolu, T. (2010). Optimal portfolio for HARA utility functions when risky assets are exponential additive processes. *Economic Notes*, **39**(1/2), pp. 65–90.

[58] Øksendal, B., and Sulem, A. (2005). *Applied Stochastic Control of Jump Diffusions*. Springer Verlag.